博士論文番号：1481202

Structural study of PHA synthase from *Chromobacterium* sp. USM2, $PhaC_{Cs}$ (クロモバクテリウム sp. USM2 由来のバイオプラスチック PHA 合成酵素 $PhaC_{Cs}$ の構造研究)

Chek Min Fey

Nara Institute of Science and Technology

Graduate School of Biological Sciences

Structural Biology Laboratory

Professor Toshio Hakoshima

Submitted on 13 SEPTEMBER 2017

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

CoA	Coenzyme A
PHA	Polyhydroxyalkanoate
PhaA	β-ketothiolase
PhaB	NADPH-dependent acetoacetyl-CoA reductase
PhaC	PHA synthase
3HA-CoA	3-hydroxyacyl Coenzyme A
3HB-CoA	3-hydroxybutyryl Coenzyme A
PHA_{SCL}	short-chain-length PHA
PHA_{MCL}	medium-chain-length PHA
$PHA_{SCL-MCL}$	SCL and MCL copolymerized PHA
$PhaC_{Cs}$-CAT	Catalytic domain of PhaC from *Chromobacterium* sp. USM2 (PDB 5XAV, this study) @ 1.48 Å
$PhaC_{Cn}$-CAT	Catalytic domain of PhaC from *Cupriavidus necator* (PDB 5HZ2, KNU, Korea) (PDB 5T6O, M.I.T, USA) @ 1.80 Å
$PhaC_{Cs}$-CAT+CoA	CoA-bound catalytic domain of PhaC from *Chromobacterium* sp. USM2 (this study) @ 3.10 Å
CAP	CAP subdomain of the catalytic domain of PhaC (residues 319-438 in $PhaC_{Cs}$)
LID region	A segment involved in conformation change (residues 327-386 in $PhaC_{Cs}$)
3HB	3-hydroxybutyrate (4 carbons Hydroxyalkanoate)
3HV	3-hydroxyvalerate (5 carbons Hydroxyalkanoate)
3HHx	3-hydroxyhexanoate (6 carbons Hydroxyalkanoate)
3HO	3-hydroxyactanoate (8 carbons Hydroxyalkanoate)
3HD	3-hydroxydecanoate (10 carbons Hydroxyalkanoate)
$PhaC_{Cs}$	Class I PHA synthase from *Chromobacterium* sp. USM2
$PhaC_{Cn}$	Class I PHA synthase from *Cupriavidus necator*
$PhaCE_{Av}$	Class III PHA synthase from *Allochromatium vinosum*

ABSTRACT

Petroleum-based (petrochemical) plastics have changed our way of life since its creation and have been utilized in almost every industrial sectors including packaging, health-care, advanced technological devices and automobile industry. On the flip side, production of petrochemical plastics releases tons of greenhouse gases, leaves undesired visual footprints on earth after its disposal, and disrupts nature ecosystems. One solution is to replace it with fully biodegradable plastics produced from renewable resources like polyhydroxyalkanoate (PHA) which exhibiting a comparable thermal stability and various industrial desirable physical properties. PHAs are synthesized by a wide-range of bacteria and halophilic archaea, as water-insoluble polyesters for carbon storage when exposed to excess carbon sources and limited amount of other nutrients in the surrounding environment. In PHA biosynthesis, PHA synthase (PhaC) is the key enzyme that polymerizes the acyl-moieties of the substrate acyl Coenzyme A (acyl-CoA) to a high-molecular weight PHA with concomitant release of CoA. Over the last 30 years, lacking the three-dimensional structure of PhaC has greatly hindered the progress in understanding the polymerization mechanism, the factors that determine their chain length and polydispersity, and the basis of their specificities.

In this study, the three-dimensional structure of the catalytic domain of class I PhaC from *Chromobacterium* sp. USM2 ($PhaC_{Cs}$-CAT) was determined at 1.48 Å resolution by X-ray crystallography. $PhaC_{Cs}$-CAT forms α/β hydrolase fold comprising α/β core and CAP subdomains. The active site consists of Cys291, Asp447 and His477 and is covered by the partly disordered CAP subdomain. Two clusters of water molecules were observed in the active site cavity. $PhaC_{Cs}$-CAT forms a

face-to-face dimer interacting by the CAP subdomain. The dimer arrangement of PhaC$_{Cs}$-CAT is distinct from the recently reported catalytic domain structures of PhaC from *Cupriavidus necator* (PhaC$_{Cn}$-CAT) which displays a relatively loosen conformation allowing a narrow path to the active site. In contrast, PhaC$_{Cs}$-CAT represents a closed form, which has no visible substrate entry pathway into the active site. In pursuit of the open form PhaC structure, co-crystals of PhaC$_{Cs}$-CAT and its inhibitor CoA was obtained and the PhaC$_{Cs}$-CAT+CoA complex structure was solved at 3.10 Å resolution. The complex structure revealed a dimer where a closed form protomer and an open form CoA-bound protomer interact through CAP subdomain. Comparison of the closed and open form structures reveals that a complete restructuring was observed in the LID region of CAP subdomain to allowed CoA entry and binding. The open form showed a clear substrate entry pathway and the possible binding pockets for acyl-moieties of the substrates was also discussed in structural context.

These three-dimensional structures of both open and closed PhaC$_{Cs}$-CAT provide valuable insights into the enzyme dynamics. Together with PhaC$_{Cn}$-CAT, the structural-based design should enable us to improve the enzymatic activity and its substrate specificity. By applying the obtained fundamental knowledge of this enzyme, industrial production and commercialization of the bioplastics could be achieved and may bring us a step closer to a sustainable tomorrow.

CHAPTER 1 INTRODUCTION

1.1 Plastics – an overview.

Plastics is a group of polymers with industrial desirable properties such as chemical- and light-resistance, durability, elasticity, flexibility to be molted, high thermostability and electricity insulation. Since the first mass production of plastics in the 1940s, plastics have become essential materials in almost every consumer products including packaging sector, health-care disposable products, advanced technological devices, and automobile industry[1,2]. There are now 20 different classes of plastics[2], with its annual plastic resin production reached 322 million tons in 2015, a 644% increase since 1976 (50 million tons)[3]. Today, plastics continue to benefit human society with expanding applications, improving the quality of life.

1.2 Downsides of plastics.

Although plastics are indispensable in today society, its non-biodegradability and production of plastics from finite resources has stirred concerns over its impacts on environment and sustainability. Some of the major concerned issues are plastics production from non-renewable resources, greenhouse gases production, waste management, human-health risk, and marine pollution.

Plastics are polymers produced from non-renewable petroleum and other additives, which will be broken down into monomer plastic resins. These resins will then manufactured into plastics with different additives for desirable characteristics[2]. It was estimated about 4% of total petroleum production were used for plastic manufacturing[4]. During petrochemical plastics manufacturing, hundreds of millions of

tons of carbon dioxide (CO_2) have been releasing into the atmosphere annually, contributing to global warming[2].

For the disposal of plastics, there are three major approaches, namely landfilling, incineration, and recycling[4]. Landfilling is the conventional waste management system, but it is challenging for countries with scarce land resources, like Japan and Denmark. The second option, incineration did reduce the requirement for space, but there are concerns that the release of hazardous substances into the atmosphere, such as dioxins, polychlorinated biphenyls and furans[4]. In a positive view, incineration recover some of the energy content in the plastic and used for electricity generation[4]. While some of the plastics can be recycled, the difficulties in collecting and sorting of the plastics has greatly suppressed the potential of plastic recycling since it is more economic to manufacture new plastic resins[4].

Also, the plastic components such as bisphenol A (BPA) and di-2-ethylhexyl phthalate (DEHP) may pose a risk to human health by acting as an endocrine-disrupting agent[1,2]. Due to the possible risks, infant bottles and spill-proof cups for toddlers manufactured using BPA has been banned in the U.S., Canada and European Union[5,6]. DEHP is the plasticizer that frequently used in polyvinyl chloride (PVC), it is not chemically bound to the plastics and there are some concerns that it might leak out easily in solution[1,7].

In the year of 2010, a total of eight million tons of mismanaged plastic waste generated by the coastal population ended up in the ocean[8]. The top 10 countries contribute to the ocean plastic debris are mainly from the developing nations, namely China, Indonesia, Philippines, Vietnam, Sri Lanka, Thailand, Egypt, Malaysia, Nigeria, and Bangladesh[8]. The ocean plastic garbage has been observed to be

accumulated in five areas and gradually formed the ocean plastic patches at the doldrums of the Atlantic, Pacific, and Indian Oceans[9]. One major concern about the plastics in the marine environment are their durability and effects on the oceanic ecosystem, marine animals and, possibly, human[10]. Weathering of plastic debris causes fragmentation of the plastic wastes into smaller particles that can be ingested by small marine animals[11]. Over 260 species of invertebrates, fish, seagulls, condor, turtles, and mammals have been documented ingesting or entangling by plastics debris[12].

Considering the usefulness of plastics, complete abandonment the use of conventional petroleum-based plastic is impractical. Instead, introduction of a biodegradable plastic into the consumer market would be the best solution.

1.3 Bio-based polymers.

Bio-based polymers can be categorized into three groups: modified natural polymers, bio-chemosynthetic polymers, and biosynthetic polymers[13] (**Table 1**). Natural polymers such as starch polymers and cellulose derivatives are modified chemically and/or physically into polymers with improved thermal and mechanical properties[14]. The bio-chemosynthetic polymers are polymerized chemically using the monomers synthesized from biological sources, such as poly(lactic acid) (PLA). For instance, lactic acid monomers are biosynthesized by microbial fermentation followed by chemical polymerization into PLA through metal catalyzed process[15]. However, not all bio-based polymers are biodegradable. For instance, crystalline PLA, cellulose derivatives, and polythioesters are not biodegradable[13,16].

Table 1. Examples of bio-based and biodegradable polymers (Adapted from Ref[27]).

Category	Processes involved	Example	Biodegradability
Modified natural polymers	i. Chemical modification of natural polymer	• Starch polymer cellulose derivatives Proteins	• Enzymatically degradable
Biochemosynthetic polymers	i. Biosynthesis of monomer ii. Chemical polymerization	• Poly(lactic acid) • Poly(butylene succinate) • Polyvinyl alcohol • Polyglycolic acid • Polythioesters	• Hydrolytically degradable except crystalline poly(lactic acid), polythioesters
Biosynthetic polymers	i. Biosynthesis of polymer by microoganisms	• Poly(3-hydroxybutyrate)	• Enzymatically degradable • Hydrolytically degradable

Biosynthetic polymers such as polyhydroxyalkanoate (PHA) is the only biodegradable polymers that completely synthesized by microoganisms[17]. It can be broke down into carbon dioxide and water under aerobic conditions and into methane under anaerobic conditions. The degradation rate of PHA is about 2 months on soil surface under tropical condition and less than 2 months in soil with more microbial exposure[13]. PHA synthesized by bacteria matches the molecular weight of conventional petrochemical plastics such as polyethylene (PE) and polypropylene (PP)[18,19]. With this character, PHA can be processed thermally into various industrial favorable forms, which could make it a favorable choice for manufacturing using current technology. Together with all the advantages, PHA is considered as one of the best candidate in replacing petroleum-based plastic.

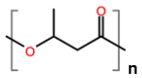
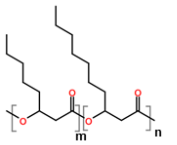
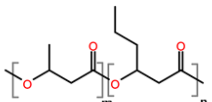
1.4 Biodegradable plastics - Polyhydroxyalkanoate (PHA).

Since the first discovery of PHA back in the year of 1926, more than 75 different genera of both Gram-positive and Gram-negative bacteria have been documented to synthesize PHA[20]. Up to now, there are more than 150 different monomeric blocks of

PHA identified[21]. Polyhydroxyalkanotes (PHAs) are biopolyesters synthesized as carbon and energy storage compounds in various bacteria under nutrient stress conditions, where there are excess carbon sources and limited amounts of other nutrients such as nitrogen, phosphate, oxygen, and sulfur in the external environment[22]. These PHAs are biosynthesized in bacterial cells as water insoluble granules up to 90% of their dry cell weight[18]. During carbon starvation, these water-insoluble PHAs will be degraded by PHA depolymerase (PhaZ) for survival[23,24]. Generally, there are three types of PHAs categorized based on the carbon numbers in the monomer building blocks: short-chain-length PHA (PHA_{SCL}), medium-chain-length PHA (PHA_{MCL}), and combination of scl and mcl ($PHA_{SCL-MCL}$) (**Table 2**).

Table 2. Types of PHA produced by bacteria from various carbon sources.

PHA_{SCL}, short-chain-length PHA; PHA_{MCL}, medium-chain-length PHA; 3HB, 3-hydroxybutyrate; 3HO, 3-hydroxyoctanote; 3HD, 3-hydroxydecanoate; 3HHx, 3-hydroxyheaxanoate.

Type of polymers	Co-polymers	Bacterial strain	Substrate
PHA_{SCL} C3-C5 carbon e.g. **P(3HB)**	 (*R*)-3HB	*Cupriavidus necator*	• Propionic acid • Pentanoic acid
PHA_{MCL} C6-C14 carbon e.g. **P(3HO-*co*-3HD)**	 (*R*)-3HO (*R*)-3HD	*Pseudomonas oleovorans*	• Gluconate
$PHA_{SCL-MCL}$ mix carbon chain e.g. **P(3HB-*co*-3HHx)**	 (*R*)-3HB (*R*)-3HHx	*Aeromonas caviae*	• Plant oils

The monomer unit in PHA_{SCL} consists of C3 to C5 carbon chain lengths such as poly(3-hydroxybutyrate) [P(3HB)]. PHA_{MCL} comprises C6 to C14 carbon chain lengths. A natural example of PHA_{MCL} is co-polymer consists of C8 3-hydroxyoctanoate (3HO) and C10 3-hydroxydecanoate (3HD) making P(3HO-*co*-3HD). The differences in monomer units leads to differences in physical properties of PHA. The copolymers categorized into mixed $PHA_{SCL-MCL}$ often possess a better thermal and physical properties compared to homopolymers like PHA_{SCL} and PHA_{MCL}. One of the commercially attractive polymer, P(3-hydroxybutyrate-*co*-3-hydroxyhexanoate) [P(3HB-*co*-3HHx)], comprising of 4 carbons 3HB and 6 carbons 3HHx possesses improved mechanical property and processability compared to P(3HB) and P(3HB-*co*-3HV)[25,26]. The copolymer P(3HB-*co*-3HHx) will become softer and more flexible with increasing fraction of 3HHx[27].

1.5 PHA biosynthesis.

To date, there are a total of 14 PHA biosynthesis pathways reported[28]. The major metabolic pathways for PHA biosynthesis are (I) the PHA_{SCL} P(3HB) biosynthetic pathway represented by *Cupriavidus necator*, (II) the fatty acid β-oxidation dependent PHA_{MCL} biosynthetic pathway represented by *Pseudomonas oleovorans*, (III) the fatty acid *de novo* biosynthetic dependent PHA_{MCL} biosynthetic pathway represented by *Pseudomonas aeruginosa*[29,30] (**Fig. 1**).

One of the simplest and extensively studied pathways is the biosynthesis pathway of P(3HB) using the model microorganism, *Cupriavidus necator* (**Fig. 1**). Metabolism of available carbon sources (sugar or fatty acid) that can be fashioned

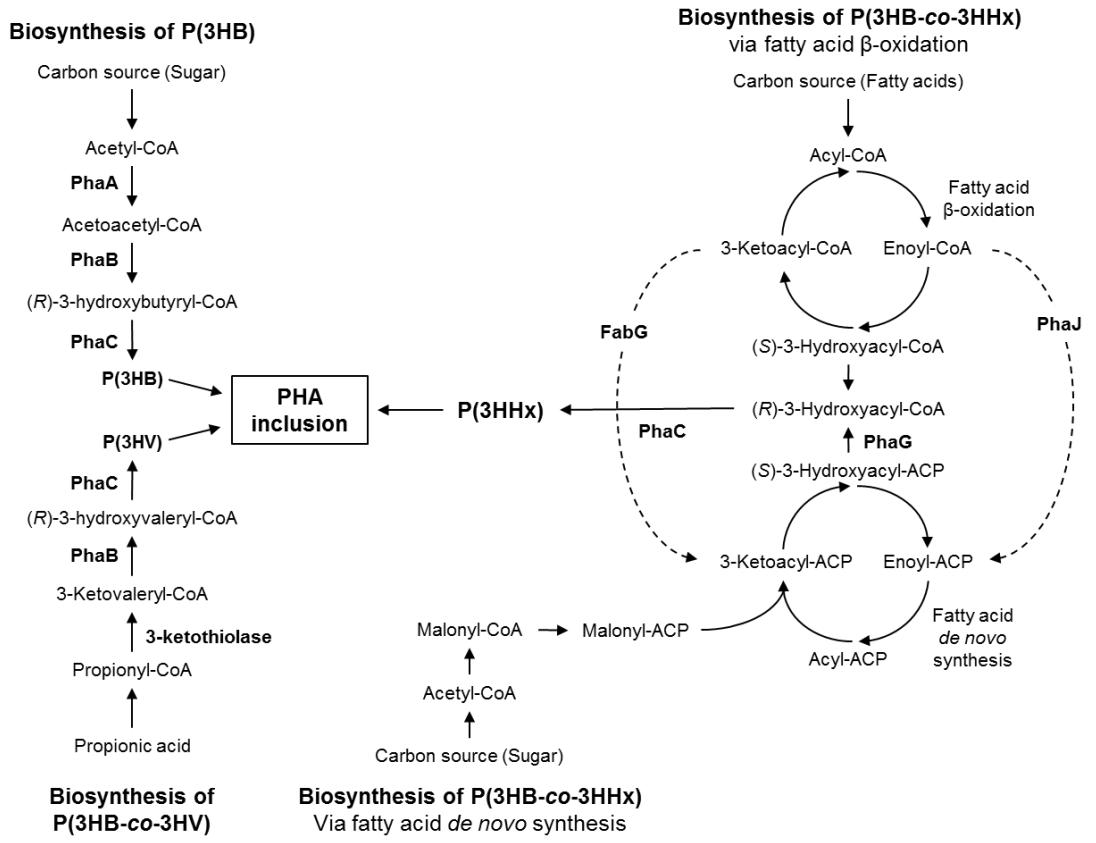


Figure 1. Major metabolic routes for PHA biosynthesis in bacteria (Adapted from Refs[29,30]). PhaA, β-ketothiolase; PhaB, NADPH depend acetoacetyl-CoA reductase; PhaC, PHA synthase; PhaG, 3-hydroxyl-ACP-CoA transferase; PhaJ, (*R*)-enoyl-CoA hydratase; FabG, 3-ketoacyl-CoA reductase.

into acyl-CoA, a precursor for PHA polymerization, is followed by the process involving a β-ketothiolase (PhaA) and NADPH-dependent acetoacetyl-CoA reductase (PhaB) for acyl-CoA synthesis and polymerization of the acyl moiety of acyl-CoA catalyzed by PHA synthase (PhaC)[29,30,31] (**Fig. 2**).

1.6 PHA synthase (PhaC).

PHA synthase (PhaC) is the key enzyme involved in polymerizing the monomeric

hydroxyalkanoate substrates without using a template[31,32,33]. The preferable substrate of PhaC is 3-hydroxyacyl-CoA with the specificity of the (*R*)-enantiomer form, such as 3-hydroxybutyryl-CoA (3HB-CoA). During polymerization, only the acyl moieties of the CoA are polymerized with a concomitant release of CoA, producing water-insoluble PHA (**Fig. 2**). PHA synthase consists of a conserved lipase box motif Gx_1Sx_2G (where x_1 and x_2 representing any amino acid residues, and Ser is replaced with Cys in PhaC)[34]. The sequences similarity with hydrolases suggest that PhaCs possess catalytic triads with the nucleophilic Cys, an acid Asp, and a general base His residues in this similar sequence order in the primary amino acid sequences.

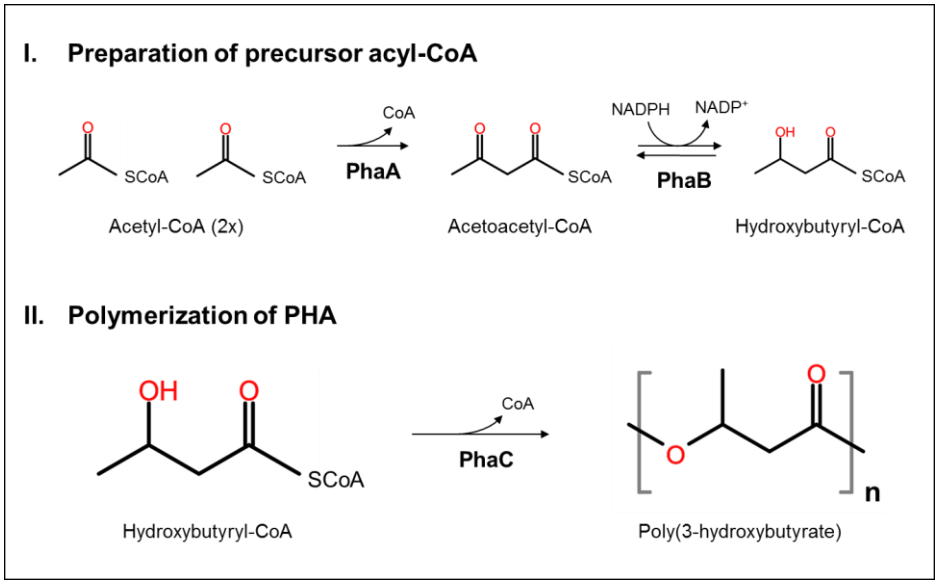


Figure 2. Schematic presentation of PHA polymerization.

PhaA, β-ketothiolase; PhaB, NADPH depend acetoacetyl-CoA reductase; PhaC, PHA synthase, respectively.

1.6.1 Classification of PhaC.

PHA synthases have been categorized into four major classes based on their primary sequences, substrate specificity, and subunit composition[31,36] (**Fig. 3**). A

paradigm for class I is PhaC from *C. necator* (old code *Ralstonia eutropha*). Together with class III synthase (e.g., in *Allochromatium vinosum*) and class IV synthase (e.g., in *Bacillus megaterium*), class I synthase has a high substrate specificity toward short-chain-length (SCL) monomers which comprise C3-C5 chain lengths. A common example of C4 SCL monomer is (*R*)-3-hydroxybutyrate (3HB). The difference between these three classes of synthases is the subunit composition where class I PHA synthase has only one PhaC subunit. In contrast, class III and class IV PHA synthases form heterodimers, comprising PhaC-PhaE and PhaC-PhaR, respectively. The representative of class II PHA synthase is that from *P. aeruginosa*, consisting of two subunits of PhaC1 and PhaC2, and is the only synthase class that favors medium-chain-length (MCL) monomeric substrates consisting C6-C14 carbon chain lengths, such as the C6 monomer 3-hydroxyhexanoate (3HHx).

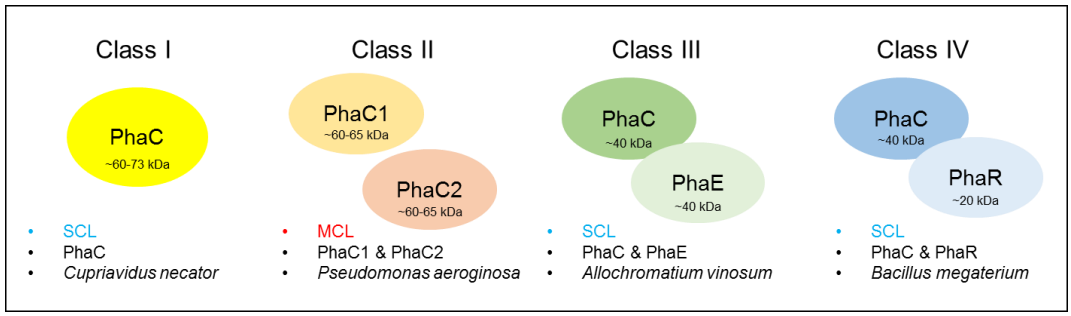


Figure 3. Classes of PHA synthases based on substrate specificity and subunit composition.

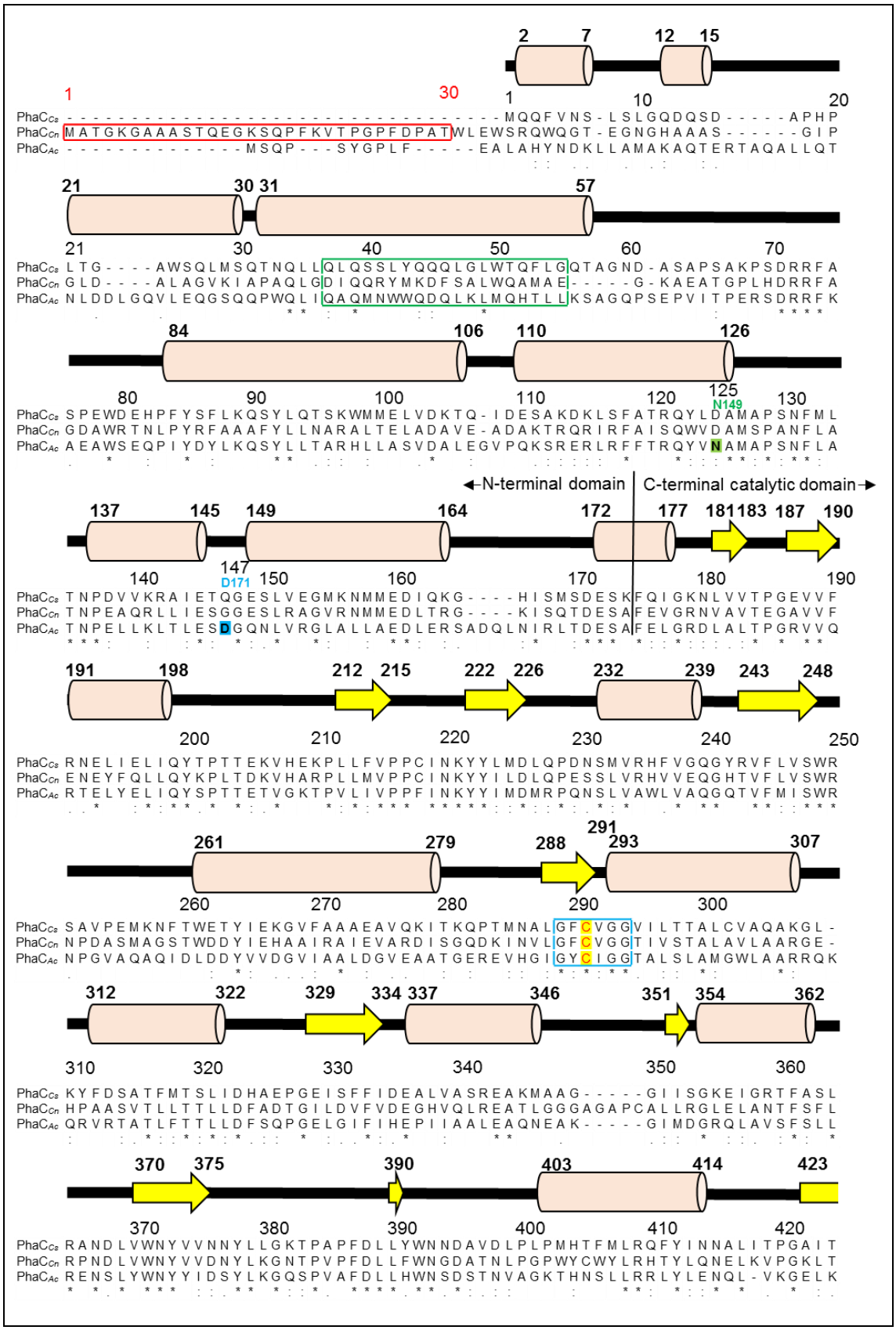
In biology, there are always exceptions. For instances, PhaC from *Aeromonas caviae* (old code *Aeromonas punctata*), with strong similarity to class I synthases, is able to catalyze the copolymer of P(3HB-*co*-3HHx)[37]. Besides, synthases PhaC1 and PhaC2 from *Pseudomonas* sp. 61-3, with strong similarity to other class II synthases,

are able to catalyze mixed copolymer of SCL 3HB with other MCL monomers[26].

1.6.2 Molecular features of PhaC.

PhaC can be categorized into N-terminal domain and C-terminal catalytic domain based on the multiple-sequences alignment and BLAST sequence-homology search. Take class I PhaC from *Chromobacterium* sp. USM2 ($PhaC_{Cs}$) as an example, its N-terminal domain (ND) comprised from residues 1 to 174 and C-terminal catalytic domain (CAT) consists from residues 175 to 567 (**Fig. 4**). In general, CAT showed high similarity to lipases (~40 kDa) and, thus, presumed as a member of α/β hydrolase superfamily[38]. In the hydrolase superfamily, all members possess an active center nucleophile that responsible for its catalytic reaction. The active site residues of hydrolases form a conserved catalytic triad comprised of a nucleophile (Cys/Ser/Asp), His and an acid (Asp/ Glu)[31,39]. The typical features of α/β hydrolase fold include 8 strands of β sheet (only second β sheet is antiparallel) and surrounded by α helices on the both sides. Typically, the crossing angle difference between the first strand and the last strand is approximately 90 degree. One conserved helix αC which forming the 'nucleophile elbow' holding the catalytic Cys is important for substrate binding[40]. Although there is no crystal structure of PhaC, several homology modeling structures of class I, II and III catalytic domain of PhaC have been performed using the crystal structures of lipases (a member of hydrolase superfamily) as templates[41-43].

The highly variability in primary sequences of the N-terminal domain of PhaC has leads to the thought that N-terminal domain is not necessary for PhaC activity[31]. For instance, first 30 amino acid residues truncated $PhaC_{Cn}$ able to display wild-type level activity in previous study[44]. Despite the exact function of N-terminal domain is


N-terminal domain
C-terminal catalytic domain
N149
D171

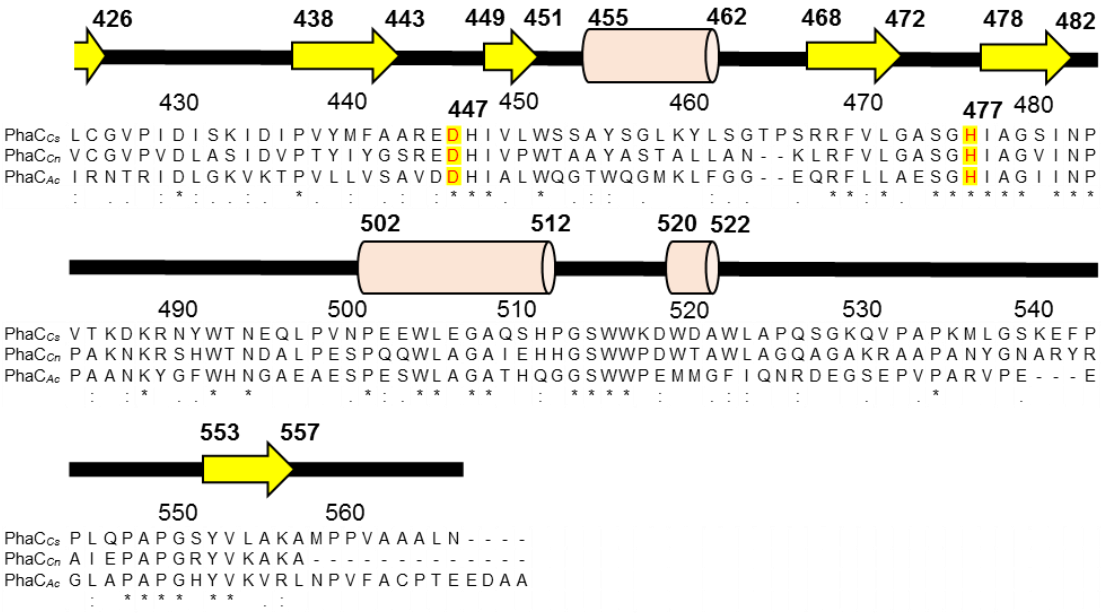


Figure 4. Secondary structure prediction of full-length PhaC$_{Cs}$ and sequences alignment with PhaC$_{Cn}$ and PhaC$_{Ac}$.

Predicted secondary structure of PhaC$_{Cs}$ using JPred4 was showed in helices and strands on top of the sequences. PhaC can be categorized into N-terminal domain and C-terminal catalytic domain. The first 30 amino acid residues (red box) of PhaC$_{Cn}$ was reported to be dispensable for its catalytic function. The predicted α-helix (D70 – E88, green box) in PhaC$_{Cn}$ (Q37 – G55 in PhaC$_{Cs}$) which is reported to be important for PhaC activity. The beneficial mutations at N-terminal domain was reported at N149 and D171 of PhaC$_{Ac}$ (corresponds to D125 & Q147 in PhaC$_{Cs}$). A conserved motif (GxCxG) similar to lipase box is showed in blue box. Catalytic triad C291, D447 and H477 of PhaC$_{Cs}$ are labeled in yellow. PhaC$_{Cs}$, PhaC from *Chromobacterium* sp. USM2; PhaC$_{Cn}$, PhaC from *Cupriavidus necator*; PhaC$_{Ac}$, PhaC from *Aeromonas caviae*; All three PhaCs are belong to class I.

still unknown and thought to be dispensable, other study has presented that truncation of first 88 amino acids or more in PhaC$_{Cn}$ diminished its catalytic activity[45]. Secondary structure prediction revealed this truncated region (1 – 88 residues in PhaC$_{Cn}$) includes a potential α-helix secondary structure comprises of Asp70 – Glu88 in PhaC from *Cupriavidus necator* (PhaC$_{Cn}$) (corresponds to Gln37 – Gly55 in PhaC$_{Cs}$, green box, **Fig. 4**)[45]. Mutational studies suggest this potential α-helix is

important in maintaining PhaC catalytic activity as well as regulating molecular weight and polydispersity of the synthesized PHA[45]. In another mutational study on the N-terminal region of PhaC from *Aeromonas caviae* ($PhaC_{Ac}$), double mutations of N149S and D171G (NSDG) displayed enhanced incorporation of 6C 3-hydroxyhexanoate monomer units into the PHA copolymer (corresponds to D125 & Q147 in $PhaC_{Cs}$**, Fig. 4**). Besides, NSDG mutant of $PhaC_{Ac}$ is able to polymerize P(3HB) homopolymer to a very high molecular weight from a relative cheaper carbon source, fructose[46]. Since PhaC is active in dimeric form, N-terminal domain also thought to stabilize the dimer form of PhaC. However, further verification and investigation are required for delineating the function of N-terminal domain.

1.6.3 Proposed catalytic mechanism of PhaC.

Catalytic triad is comprised of a group of amino acid residues working together in a charge relay system at the core active site of most hydrolases, such as lipases and proteases. Interestingly, PhaCs also possess a conserved motif similar to the lipase box (GxSxG) where the Ser is substituted by Cys in PhaC (green box, **Fig. 4**)[47]. A conserved catalytic triad in PhaC comprised of a nucleophile Cys, a general base His, and an acid Asp was identified and the importance of the catalytic triad was experimental confirmed by mutational studies[38].

The first proposed catalytic mechanism of PhaC was based on fatty acid synthases (FAS)[35]. In this model, two thiol groups involved in the covalent catalysis, which might be the thiol groups from the catalytic Cys residues from the dimeric PhaC. First thiol group acts as the loading site for substrate acyl-CoA while the second thiol group acts as the priming and elongation site. However, the sequence

similarity between PhaC and FAS is low and no second thiol residue identified in PhaC. These observations have cast doubts on the validity of this proposed mechanism as there might be structural constraint to bring both cysteine together[35].

Since the conserved lipase box was observed in PhaC, an alternative mechanism was proposed based on the catalytic mechanism of lipase[38,48]. In this model, the polymerization of PHA could be catalyzed by a single thiol group from catalytic cysteine. The catalytic histidine functions as a general base to activate the cysteine to attack thioester of the first substrate acyl-CoA and form Cys-acyl intermediate[38]. Then, in the elongation step, Asp activates the hydroxyl group from the second substrate acyl-CoA. Then, activated hydroxyl group will attack the Cys-acyl intermediate to form 2 acyl monomers product that is covalently or non-covalently bound to the PHA synthase[35].

The polymerization mechanism of both models is highly similar, but the number of thiol groups involved is different. Hence, structural information showing distance of the active sites from each protomer in the dimeric PhaC is required to have a better understanding in its catalytic mechanism.

1.7 PHA synthases from *Chromobacterium* sp. USM2 (PhaC$_{Cs}$).

To produce PHA$_{SCL-MCL}$ with better thermal and physical properties, a lot of efforts have been placed in PhaC engineering by random mutagenesis to broaden its substrate specificities[25,49]. Alternatively, isolation and screening of natural evolved synthases with higher performances and broader substrate specificity is still a common practice, especially in megadiverse countries like Malaysia. A class I PHA synthase gene (*phaC$_{Cs}$*) from *Chromobacterium* sp. USM2 which was newly isolated

in an island named Langkawi near the northern Peninsular of Malaysia[50]. *Chromobacterium* sp. USM2 is Gram-negative rod-shaped bacterium and deposited in Japanese Culture Collection with the accession number of JCM 15051.

Interestingly, $PhaC_{Cs}$ produces P(3HB-*co*-3HHx) copolymer and poly(3-hydroxybutyrate-*co*-3-hydroxyvalerate-*co*-3-hydroxyhexanoate) [P(3HB-*co*-3HV-*co*-3HHx)] terpolymer, which are ideal materials for use in industrial products due to their softness and flexibility[50,51]. Moreover, an *E. coli* transformant harboring $phaC_{Cs}$ was shown to accumulate P(3HB) up to 76 weight percent within 24 hours of cultivation. Furthermore, Strep2-tagged $PhaC_{Cs}$ showed a specific activity of 238 U/mg, which is almost 5-fold higher than the activity of PhaC from *Cupriavidus necator*/ *Ralstonia eutropha* ($PhaC_{Cn}$), a commonly used PHA synthases in previous studies[52].

1.8 Current bottlenecks in PHA development.

The reality of replacing conventional petrochemical polymers is possible with the recent advancement in biotechnology and growing public awareness. However, there are obstacles in commercializing PHA due to its performance and relative higher production cost. Over the past 30 years or so, attempts to obtain structural information of PHA synthase through X-ray crystallography was unsuccessful. The hitherto lack of a three-dimensional structure of PhaC has limited our understanding of the polymerization mechanisms involved, the factors that determine chain length and polydispersity, and the molecular basis of the enzyme specificity. An understanding of the mechanisms based on three-dimensional structures should play an important role in determining the feasibility of these materials for industrial use.

CHAPTER 2 EXPERIMENTAL PROCEDURES

2.1 Cloning, expression and purification. The gene of full-length PHA synthase *phaC$_{Cs}$* (1704 nucleotides) and the catalytic domain PhaC$_{Cs}$-CAT (1179 nucleotides) from *Chromobacterium* sp. USM2 (JCM15051, RIKEN BRC) were cloned from the bacterial genome and inserted into His-fusion vector pET47b [+] (Novagen). Catalytic triad mutant plasmid of D447N was synthesized by PCR site-directed mutagenesis using wild-type plasmid as template. The integrity of the coding region was verified by DNA sequencing. The plasmid was transformed into *Escherichia coli* Rosetta 2 (DE3) (Novagen) and cells were grown in LB medium supplemented with 50 µg/ml Kanamycin and 35 µg/ml Chloramphenicol at 37˚C until the OD_{600nm} reached a value of 0.6. Protein expression was induced by the addition of isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 100 µM. The culture was induced for another 20 hours at 20˚C and cells were harvested by centrifugation. The cells were then suspended in 2X PBS (phosphate-buffered saline) with 3 mM β-mercaptoethanol (β-ME) and disrupted by sonication in an ice bath. The soluble fraction was separated by ultracentrifugation and then loaded onto a Ni-NTA agarose column (QIAGEN). The column was washed with buffer containing 20 mM Tris-Cl (pH 8.0), 100 mM NaCl, 3 mM β-ME and 10 mM Imidazole, and target protein was eluted using the same buffer containing 250 mM Imidazole in lieu of 10 mM Imidazole. The column effluent was loaded onto a HiTrap Q anion-exchange column (GE Healthcare) and eluted using a gradient of 0 to 500 mM NaCl in buffer containing 20 mM Tris-Cl (pH 8.5) and 3 mM β ME. The eluted products were pooled and digested overnight at 4˚C using HRV 3C protease

(Merck). The digested product was loaded onto a Ni-NTA agarose column to remove the His-tag and co-purified products. Flow-through from the column was collected and concentrated by centrifugation using an Amicon Ultra 30,000 molecular weight cut-off filter (Merck Millipore). The concentrated product was finally passed through a Superdex 200 gel filtration column (GE Healthcare) using buffer comprising 10 mM Tris-Cl (pH 8.0), 100 mM NaCl and 3 mM β-ME. Peak fractions were collected and concentrated to about 50 mg/ml. SDS-PAGE of the protein samples gave one major band corresponding to ~63 kDa (full-length PhaC$_{Cs}$) and ~43 kDa (PhaC$_{Cs}$-CAT). Analysis of the sample using matrix-assisted laser desorption/ ionization time-of-flight mass spectrometry (MALDI-TOF MS; Bruker Daltonics) confirmed that the target proteins were successfully purified without degradation. Two additional vector-derived residues (Gly-Pro) were identified at the N-terminus. The catalytic mutant D447N of PhaC$_{Cs}$-CAT were purified using the same procedures as wild-type PhaC$_{Cs}$-CAT. Samples were frozen in liquid nitrogen and stored at -80˚C until use.

2.1.1 Expression of Selenomethionine-labeled full-length PhaC$_{Cs}$.

Selenomethionine (SeMet)-labeled full-length PhaC$_{Cs}$ was expressed in M9 medium (**Table 3**) containing SeMet under conditions that inhibited the methionine biosynthetic pathway. Firstly, fresh single colony of *E. coli* containing plasmid of PhaC$_{Cs}$ was grow overnight in LB medium at 37°C as primary seed culture with appropriate antibiotic. Then, the cell pellet was harvest and washed with 1X PBS to remove trace of LB medium. The cell pellet was resuspended in 100 ml of M9 media and grow at 37°C overnight as secondary seed culture. The next day, secondary seed culture was subcultured into 2 L M9 media and grow at 37°C till culture's OD$_{600nm}$

reached 0.30-0.35. The powders of 6 amino acids (200 mg Lys, 200 mg Phe, 200 mg Thr, 100 mg Ile, 100 mg Leu, and 100 mg Val) were added into the 2L M9 media and incubation continued till culture's OD_{600nm} reached 0.5. Next, 120 mg of L-SeMet added into the 2 L M9 media and further incubates for another 30 minutes. Then, the cultures were ice-bath for 15 minutes and 1 mM of IPTG was added into the culture for protein expression at 18°C for 24 hours. The cell cultures were harvest as usual. The procedures employed for protein purification were the same as described above, except that dithiothreitol (5 mM) was used in lieu of β-ME. Analysis using MALDI-TOF MS confirmed that all twenty-one methionine residues in $PhaC_{Cs}$ were successfully replaced with SeMet.

Table 3. M9 medium for expression of selenomethionine substituted proteins in common *E. coli*.

M9 medium	1 L	Sterilization
Na_2HPO_4	6.9 g	Autoclave
KH_2PO_4	3.0 g	
NH_4Cl	1.0 g	
NaCl	0.5 g	
D-glucose	4.0 g	Filtering
$MgSO_4$	250 mg	
$CaCl_2$	10 mg	
$FeSO_4$	4 mg	
KAO & MICHAYLUK Vitamin Solution (X100, SIGMA)	10 ml	

2.2 Crystallization of $PhaC_{Cs}$-CAT. Preliminary crystallization screening of $PhaC_{Cs}$ was performed using the vapor-diffusion method at both 4˚C and 20˚C using commercially available screening kits (Hampton Research and QIAGEN). Protein

(0.3 mM) was mixed at 1:1 ratio with the reservoir solution. Crystals of PhaC$_{Cs}$ were observed after one year in equilibrium against mother liquor containing 0.2 M lithium sulfate, 0.1 M Bis-Tris (pH 6.5) and 25% PEG3350 (Hampton Research) at 20˚C (**Fig. 5a**). Crystals were cryoprotected by the addition of 20% glycerol and flash-cooled using liquid nitrogen. However, the production of these tiny hexagonal crystals (<100 μm) could not be repeated using the original kit reagent or self-made mother liquor. In efforts to reproduce the crystals, full-length PhaC$_{Cs}$ was successfully crystallized in the presence of α-chymotrypsin protease at a ratio of 1 μg protease to 600 μg PhaC$_{Cs}$[53] (**Fig. 5b**). Crystals of α-chymotrypsin-digested PhaC$_{Cs}$ were grown after four days in lieu of one year under the same conditions. The crystals were cryoprotected in 0.1 M Bis-Tris (pH 6.5), 0.2 M lithium sulfate, 15% PEG3350 and 33% ethylene glycol and flash-cooled using liquid nitrogen. The crystals of α-chymotrypsin-treated SeMet-labeled PhaC$_{Cs}$ were optimized by homo- and hetero-seeding under the same conditions as described above (**Fig. 5c**). Crystals of PhaC$_{Cs}$-CAT were grown in 0.2 M ammonium sulfate and 15% PEG4000 at 4˚C after 20 days. The crystals were optimized by streak-seeding in 50 mM Bis-Tris (pH 5.5), 60 mM ammonium sulfate and 5% PEG4000 (**Fig. 5d**). The crystals were cryoprotected in 33% ethylene glycol. The determined structure was deposited as 5XAV in PDB.

2.3 Co-crystallization of PhaC$_{Cs}$-CAT (D447N) and CoA complex. Preliminary co-crystallization screening was performed using vapor-diffusion method at both 4˚C and 20˚C using commercially available screening kits (Hampton Research and QIAGEN). To prepare the stable PhaC$_{Cs}$-CAT+CoA complex, the PhaC$_{Cs}$-CAT

(D447N) (0.3 mM) was mixed with CoA (12 mM) in equilibration buffer containing 10 mM Tris-Cl (pH 8.0), 100 mM NaCl, 3 mM β-ME. The mixture was incubated at 30°C for 30 minutes prior to crystallization set-up. The first hit was a cluster of crystals observed in equilibrium against mother liquor containing 0.2 M sodium acetate and 20% PEG3350 (QIAGEN) at 4˚C. The optimized crystals grown in 5 days using streak-seeding method against a self-made mother liquor containing 0.1 M Tris-Cl (pH 7.0), 0.015 M sodium acetate, and 9% PEG3350 at 4˚C (**Fig. 5e**). The crystals were cryoprotected in the same mother liquor with 38-40% PEG3350.

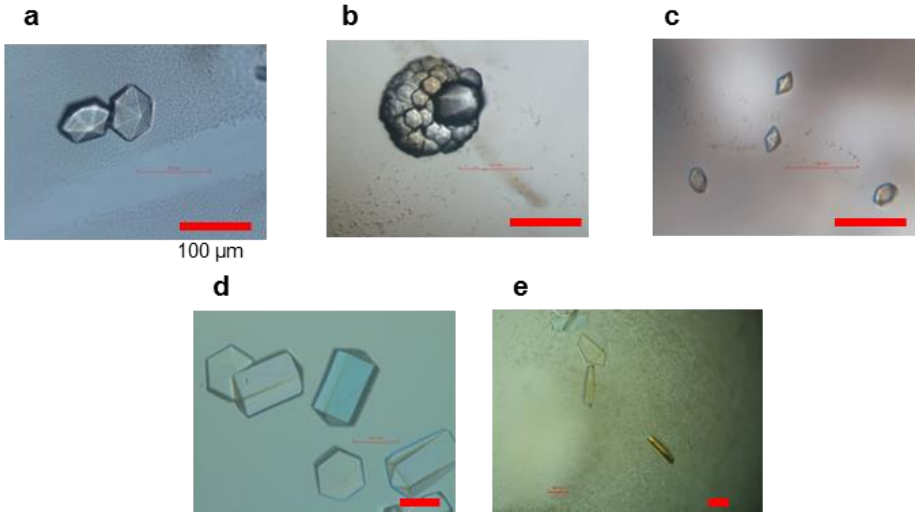


Figure 5. Obtained crystals of PhaC$_{Cs}$-CAT and PhaC$_{Cs}$-CAT+CoA.

(**a**) Hexagonal crystals of full-length PhaC$_{Cs}$ grown after 8 – 12 months in a sitting drop set up equilibrated against INDEX (75): 0.1 M Bis-Tris (pH6.5), 0.2 M Lithium sulfate, and 25% PEG3350.
(**b**) Crystal clusters of α-chymotrypsin-digested PhaC$_{Cs}$ in a hanging drop against 0.1 M Bis-Tris (pH6.5), 0.2 M Lithium sulfate, and 20% PEG4000.
(**c**) α-chymotrypsin-digested Se-Met labeled PhaC$_{Cs}$ grown in 0.1 M Bis-Tris (pH6.3), 0.2 M Lithium sulfate, and 15% PEG3350 by seeding.
(**d**) Hexagonal crystals of PhaC$_{Cs}$ (175-567) grown in 50 mM Bis-Tris (pH5.5), 0.06 M Ammonium sulfate, and 5% PEG4000 by seeding.
(**e**) Complex crystals of PhaC$_{Cs}$ (175-567) D447N and CoA grown in 100 mM Tris-Cl (pH 7.0), 1.5 mM sodium acetate, and 9% PEG3350 by seeding.

2.4 X-ray data collection, phasing and refinement. All X-ray data were

collected at SPring-8, Harima, Japan. During X-ray beam exposure, crystals were flash-cooled and maintained at 100 K using a nitrogen stream. Detailed statistics of the structure determination are shown in **Table 4**. All diffraction data were indexed and merged using the DENZO and SCALEPACK programs included in HKL2000[54]. The phase problem of PhaC$_{Cs}$-CAT was solved by single-wavelength anomalous dispersion (SAD) using an α-chymotrypsin-digested SeMet-labeled PhaC$_{Cs}$ crystal. Determination of the selenium positions, phasing, density modification, and initial modeling of the SeMet-labeled PhaC$_{Cs}$ structure were performed using the autoSHARP program suite[55] comprising the heavy atom search program ShelxC/D[56], the phasing program SHARP47, and density modification program SOLOMON[57]. Buccaneer[58] and ARP/wARP[59] were then employed to build the model automatically. Subsequent manual model building and refinement were performed using refmac5[60], phenix.refine[61] and Coot[62]. Other structures were determined by molecular replacement using the Phaser program[63]. Finally, the model structure of free form PhaC$_{Cs}$-CAT were further refined using refmac5, phenix.refine and Coot. Structure of PhaC$_{Cs}$-CAT+CoA complex was solved by molecular replacement using the structure of free form PhaC$_{Cs}$-CAT (PDB 5XAV) as the reference model. Superposition of the PhaCs and lipases was performed using the program LSQKAB[64]. Illustrations were prepared using the program PyMOL (DeLano Scientific). Fitting of acyl moieties into the possible binding pocket was manually modeled using Coot[62].

2.5 Analytical centrifugation (AUC). Sedimentation velocity ultracentrifugation experiments were performed using a Beckman Coulter Optima XLA analytical ultracentrifuge at 20˚C. Full-length, PhaC$_{Cs}$ (calc. 63.4 kDa) samples were dissolved

Table 4. Crystallographic statistics of free form $PhaC_{Cs}$-CAT and complex $PhaC_{Cs}$-CAT+CoA.

Crystallographic Analysis Statistics			
Crystal form	Native (Free form) $PhaC_{Cs}$-CAT	D447N (CoA-bound complex) $PhaC_{Cs}$-CAT	SeMet (SAD)[a] (α-chymotrypsin treated) Full-length $PhaC_{Cs}$
Space group	$P6_1$	$C2$	$P6_2$
Unit cell			
a , b , c (Å)	117.33, 117.33, 105.91	169.80, 61.85, 82.57	117.15, 117.15, 52.69
α, β, γ (°)	90, 90, 120	90.00, 106.73, 90.00	90, 90, 120
SPring-8 Beamline	BL44XU	BL41XU	BL44XU
Wavelength (Å)	0.90000	0.97914	0.97906 (Se peak)
Resolution range[b] (Å)	50.00 - 1.48 (1.53 - 1.48)	50.00 – 3.10 (3.15 – 3.10)	50.00 - 2.50 (2.54 - 2.50)
Reflections			
Oscillation range (°)	360	360	360
Measured/Unique	1535,125/137,720	47983/15234	304,374/14,563
Multiplicity	11.2 (11.2)	3.3 (2.9)	20.9 (22.5)
Mosaicity (°)	0.17 – 0.34	0.334 –.611	0.44 – 0.73
$I/s(I)$	58.6 (6.2)	15.2 (1.7)	73.5 (15.5)
R_{merge}(%)	7.5 (64.4)	7.9 (51.3)	12.2 (76.4)
Completeness (%)	99.8 (100.0)	96.5 (87.5)	100.0 (100.0)
Number of heavy atoms[c]	-	-	5
FOM (acentric/centric)[d]	-	-	0.365/0.161
Refinement Statistics			
R_{work} / R_{free} (%)[e]	12.11/15.74	22.78/26.30	
Number of atoms	6968	5918	
Protein molecules	2 (754 residues)	2 (741 residues)	
Water molecules	1050	14	
Ligand (Coenzyme A)	-	48	
Average B-factors (Å^2)			
Protein	20.1	99.50	
Water molecules	33.5	82.63	
Ligand (Coenzyme A)	-	138.08	
R.m.s.d. from ideal values			
Bonds (Å)/ angles (°)	0.019/1.603	0.004/0.66	
Ramachandran plot (%)			
Favored	97.0	93.45	
Allowed	2.7	6.41	
Outliers	0.4	0.14	

[a] The SeMet derivative was treated with α-chymotrypsin.
[b] Values in parentheses are for the highest-resolution shell.
[c] Heavy atoms were searched with the program, SHELX C/D.
[d] The figure of merit (FOM) was calculated with the program, SHARP/autoSHARP.
[e] R_{free} was calculated on a random 5 % reflections of the data.

in 10 mM Tris-Cl (pH 8.0), 100 mM NaCl and 3 mM β-ME at a concentration of 5 µM in the absence or presence of 50 µM DL-3HB-CoA and centrifuged at 20,000 rpm. Similarly, PhaC_{Cs}-CAT (calc. 43.8 kDa) samples were prepared at a concentration of 5 or 20 µM in the absence of DL-3HB-CoA. The raw results were analyzed using the program SEDFIT.

2.6 Size exclusion chromatography (SEC). Analytical gel filtration was performed with protein samples in the absence or presence of substrate by loading onto a Superdex 200 (10/30) gel filtration column in buffer containing 10 mM Tris-HCl (pH 8.0), 100 mM NaCl and 3 mM β-ME at 4˚C. A final concentration of 30 µM protein in the absence or presence of 0.5 mM DL-3HBCoA was loaded and its elution profile was compared with the standard elution profile. The molecular weights of the eluted peaks were calculated and compared with monomeric standard molecular weights.

2.7 *in vitro* enzymatic assay. PHA synthase activity assays were performed by measuring the CoA released from the substrates during the polymerization process[52]. The activities of individual mutants of PhaC_{Cs}-C218A, C218S, T310A and T310S were performed at 30˚C in a final volume of 360 µl reaction mixture containing 100 mM Tris-Cl buffer (pH 7.5), 1 mg/ml BSA, 0.6 mM DL-3HB-CoA (Sigma Aldrich), and 1 mM 5,5-dithio-bis-(2-nitrobenzoic acid). Following the addition of enzymes (7.5-30 nM) into the reaction mixture, measurement of CoA release was taken at various time points by measuring the absorbance at 412 nm using a UV spectrophotometer (Shimadzu Spectrophotometer UV-1800).

CHAPTER 3 RESULTS

3.1 Overall structure of the catalytic domain of PhaC$_{Cs}$ in the free form.

PHA synthase (PhaC) consists of two domains, the N-terminal region and C-terminal catalytic domain (**Fig. 6a**). The full-length PhaC consists of a protease-sensitive region between the N-terminal region and the C-terminal catalytic domain. The flexibility of this region might hinder crystallization of its full-length protein. However, the C-terminal catalytic domain grown well-diffracted crystals, probably due to its high structural stability. The crystal structure of the catalytic domain (residues 175-567) of PhaC$_{Cs}$ (PhaC$_{Cs}$-CAT) was determined by a single-wavelength anomalous diffraction (SAD) method (1.48 Å). The crystallographic statistic is shown in **Table 4**.

The crystal structure of PhaC$_{Cs}$-CAT reveals a globular structure adopting α/β hydrolase fold and consists of α/β core and CAP subdomains (**Fig. 6b-d**). The core subdomain forms 13 β-strands (β1-β13) and comprises the central eleven-stranded β-sheet, which contains 9 parallel and 2 antiparallel β-strands with strand connectivity β13-β1-β2-β5-β4-β6-β7-β8-β9-β10-β11, and seven α-helices (α1-α7) located at both sides of the central β-sheet: with α1, α5 and α7-helices on one side, and α2, α3, α4 and α6 on the other side (**Fig. 6c, d**). In addition to these α-helices and β-strands, eight short 3_{10}-helices (η1-η3 in the core and ηA-ηD in the CAP subdomains) were observed. Besides the central β-sheet, the core subdomain forms an additional two-stranded antiparallel β-sheet (β3 and β12) which is located beneath the core subdomain (**Fig. 6b, d**). As seen in the α/β hydrolase fold, the central β-sheet of PhaC$_{Cs}$-CAT displays a left-handed super-helical twist with the first β1 strand crossing

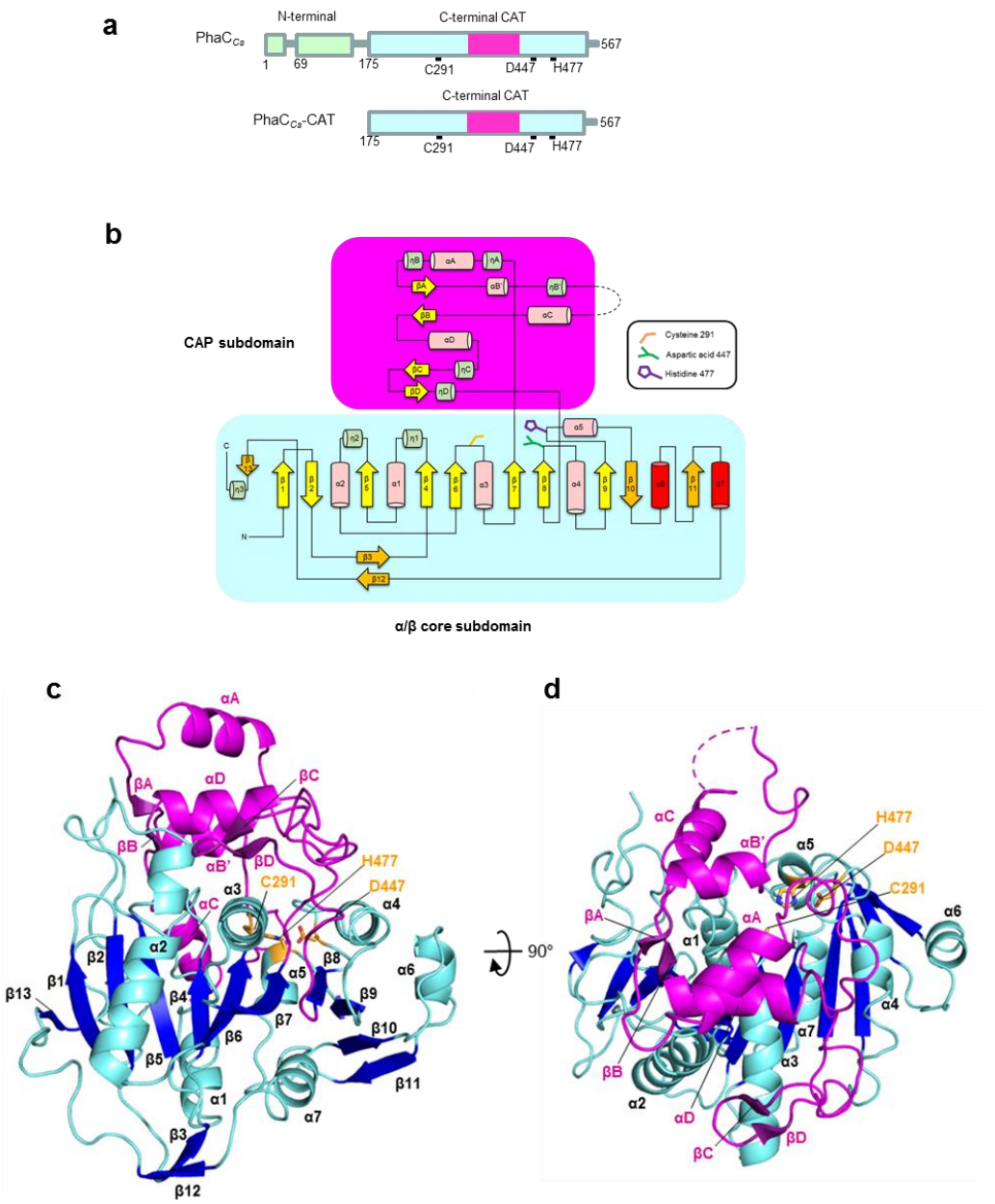


Figure 6. Structure of PhaC$_{Cs}$-CAT.

(**a**) Domain organization of PhaC$_{Cs}$. PhaC$_{Cs}$ consists of the N-terminal domain and the C-terminal catalytic (CAT) domain.

(**b**) Schematic presentation of the secondary structure topology of PhaC$_{Cs}$-CAT (mol B). The catalytic domain contains the CAP and core subdomains. The other strands (orange) and helices (red) shown represent additional secondary structures observed when comparing the canonical α/β hydrolase fold.

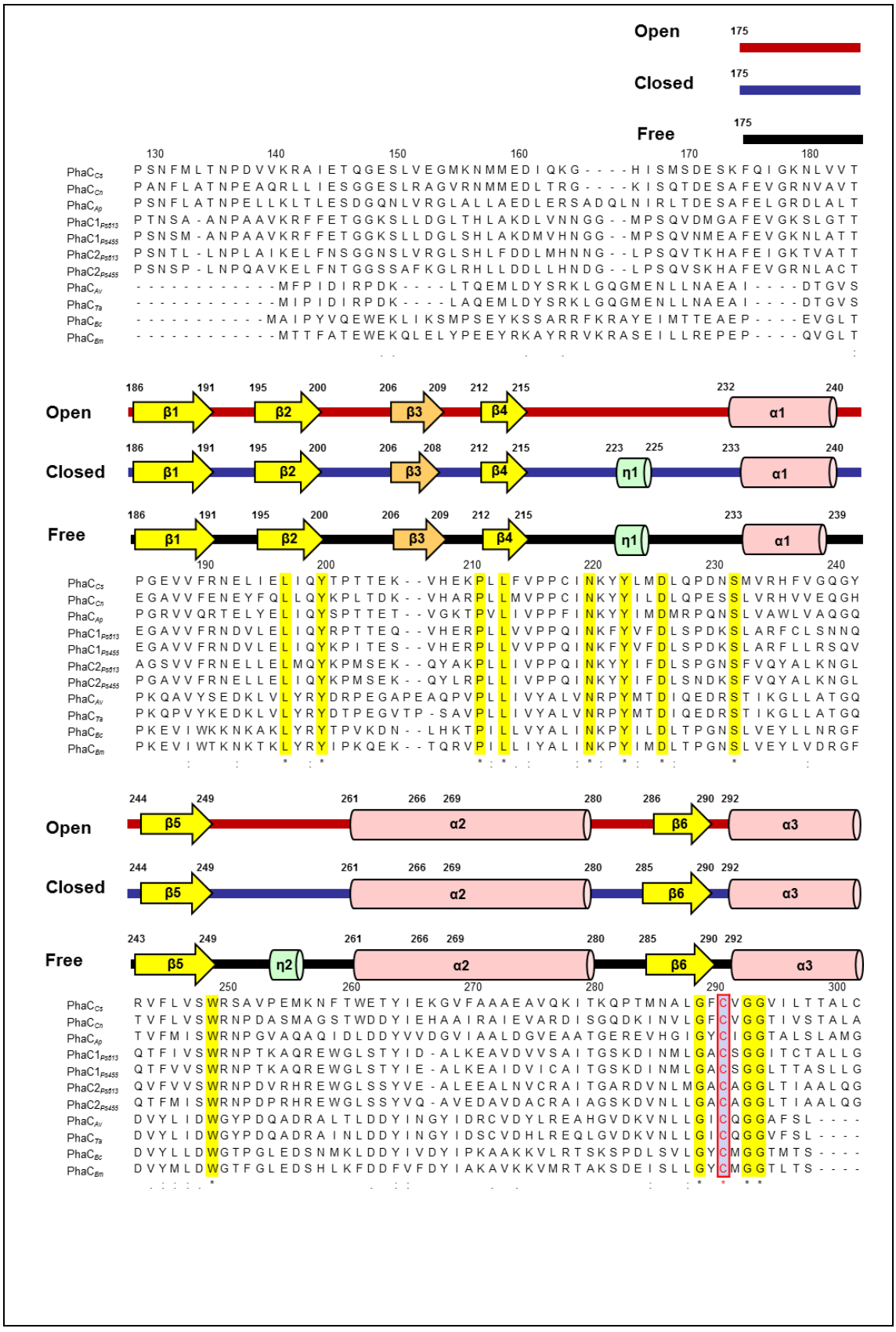
(**c**) A side view of PhaC$_{Cs}$-CAT. The catalytic domain of PhaC$_{Cs}$ comprises the α/β core subdomain (residues 175–318, 439–562 in cyan) and the CAP subdomain (residues 319–438 in violet). The side chains of the catalytic triad (Cys291, His477 and Asp447) are shown in orange.

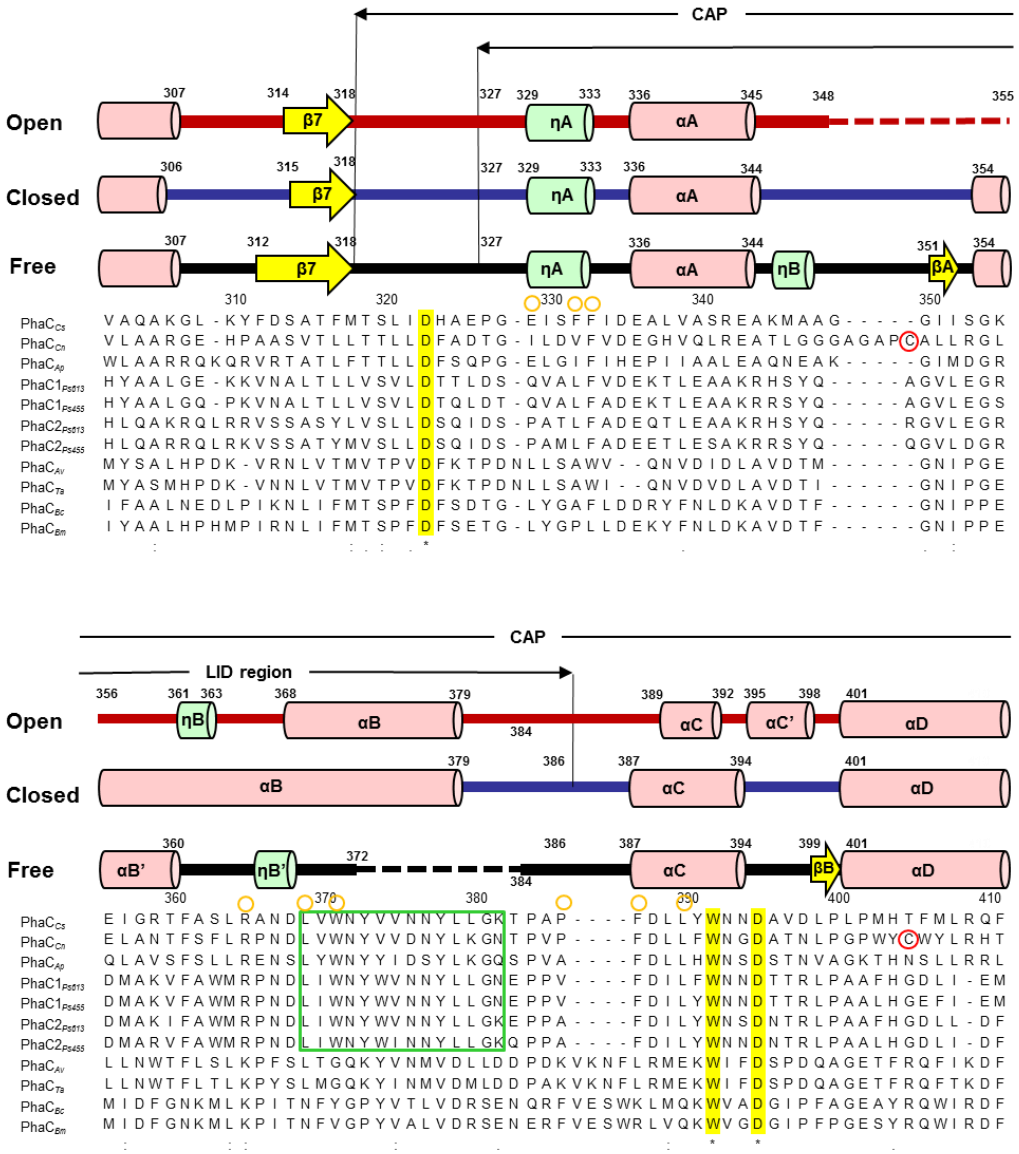
(**d**) As in **c**, but a top view of PhaC$_{Cs}$-CAT. The active site is covered by the CAP subdomain.

the eighth β-strand (β9) at ~90˚ and the last strand (β11) at ~180˚. In the core subdomain, Class III and IV synthases lack most of the C-terminal additional region following β10 strand (**Fig. 7**).

The CAP subdomain (residues 319–438) is protruding from the core α/β fold (residues 186–318 and residues 439–562) through β7 strand and connecting back to β8 strand (**Fig. 6c**). There are four α-helices (αA, αB', αC and αD) and four β-strands (βA-βD) forming the CAP subdomain with an 11-residue missing segment (373-YVVNNYLLGKT-383 in chain B) located at the random coil between ηB' and αC helices. In the structure, the cavity containing the catalytic triad was completely covered by the CAP subdomain (**Fig. 6d**).

PHA synthases have previously been described by comparison with α/β hydrolase of the lipase family[38]. The current structure of PhaC$_{Cs}$-CAT, however, reveals several unique structural characteristics, which are absent in the canonical α/β hydrolase fold of lipases constructed by the α/β core comprising eight β-strands and six α-helices. Structural comparison with mammalian lipases[65,66] showed differences in their architecture (**Fig. 8**). The major difference in the core subdomains is due to the additional presence of five β-strands (β3, β10, β11, β12, and β13) and two α-helices (α6 and α7) in PhaC$_{Cs}$-CAT. An additional major deviation is found in the helical packing of the CAP subdomain: The N-terminal 3_{10}-helix (ηA) and the C-terminal two helices (αC and αD) adopt a comparable orientation and position with that of the corresponding helices of lipases, while the other helices show no similarity due to conformational rearrangement.


Open
Closed
Free
β1
β2
β3
β4
α1
η1
β5
α2
β6
α3
η2
PhaC$_{Cs}$
PhaC$_{Cn}$
PhaC$_{Ap}$
PhaC1$_{Ps613}$
PhaC1$_{Ps455}$
PhaC2$_{Ps613}$
PhaC2$_{Ps455}$
PhaC$_{Av}$
PhaC$_{Ta}$
PhaC$_{Bc}$
PhaC$_{Bm}$



continued

The abbreviations are as follows;

Class I	PhaC$_{Cs}$:	*Chromobacterium* sp. USM2 (ADL70203)
	PhaC$_{Cn}$:	*Cupriavidus necator* (AAW65074)
	PhaC$_{Ap}$:	*Aeromonas punctata* (BAA21815)
Class II	PhaC1$_{Ps61\text{-}3}$:	*Pseudomonas* sp. 61-3 (BAA36200)
	PhaC2$_{Ps61\text{-}3}$:	*Pseudomonas* sp. 61-3 (BAA36202)
	PhaC1$_{PsUSM4\text{-}55}$:	*Pseudomonas* sp. USM4-55 (ABX64434)
	PhaC2$_{PsUSM4\text{-}55}$:	*Pseudomonas* sp. USM4-55 (ABX64435)
Class III	PhaC$_{Av}$:	*Allochromatium vinosum* DSM180 (BAE20055)
	PhaC$_{Tv}$:	*Thiocystis violascens* DSM198 (AFL75311)
Class IV	PhaC$_{Bc}$:	*Bacillus cereus* (BAI68395)
	PhaC$_{Bm}$:	*Bacillus megaterium* (AAD05260)

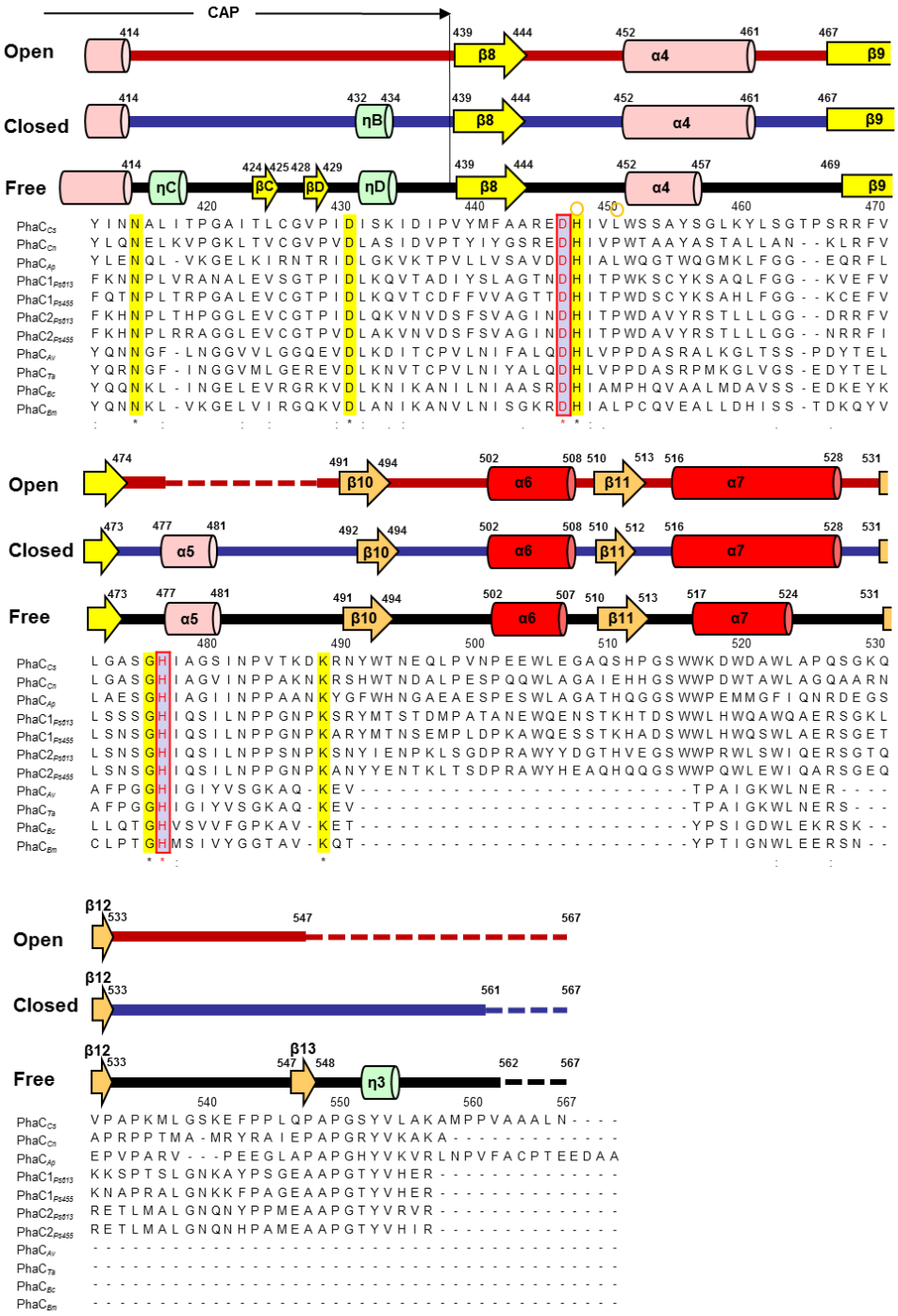


Figure 7. Alignment of amino acid sequences of PhaC catalytic domains.

The secondary structure elements found in the structures of PhaC$_{Cs}$-CAT (Free) and PhaC$_{Cs}$-CAT+CoA (Open mol B & Closed mol A) are shown at the top of the alignment with α-helices (pink cylinders), β-strands (yellow arrows), loops (bold lines) and missing loops (broken lines). Glu329, Phe332, Phe333, Arg365, His448 and Val450 of PhaC$_{Cs}$-CAT are marked with orange circles at the top. Two Cys residues (Cys382 and Cys438 of PhaC$_{Cn}$-CAT) forming a disulfide bond are marked by red circles. Part (Leu402–Asn415 of PhaC$_{Cn}$-CAT) of the LID region forming α4 helix in the structure is conserved in members of Class I and II synthases. These sequences are marked by circles within a green box.

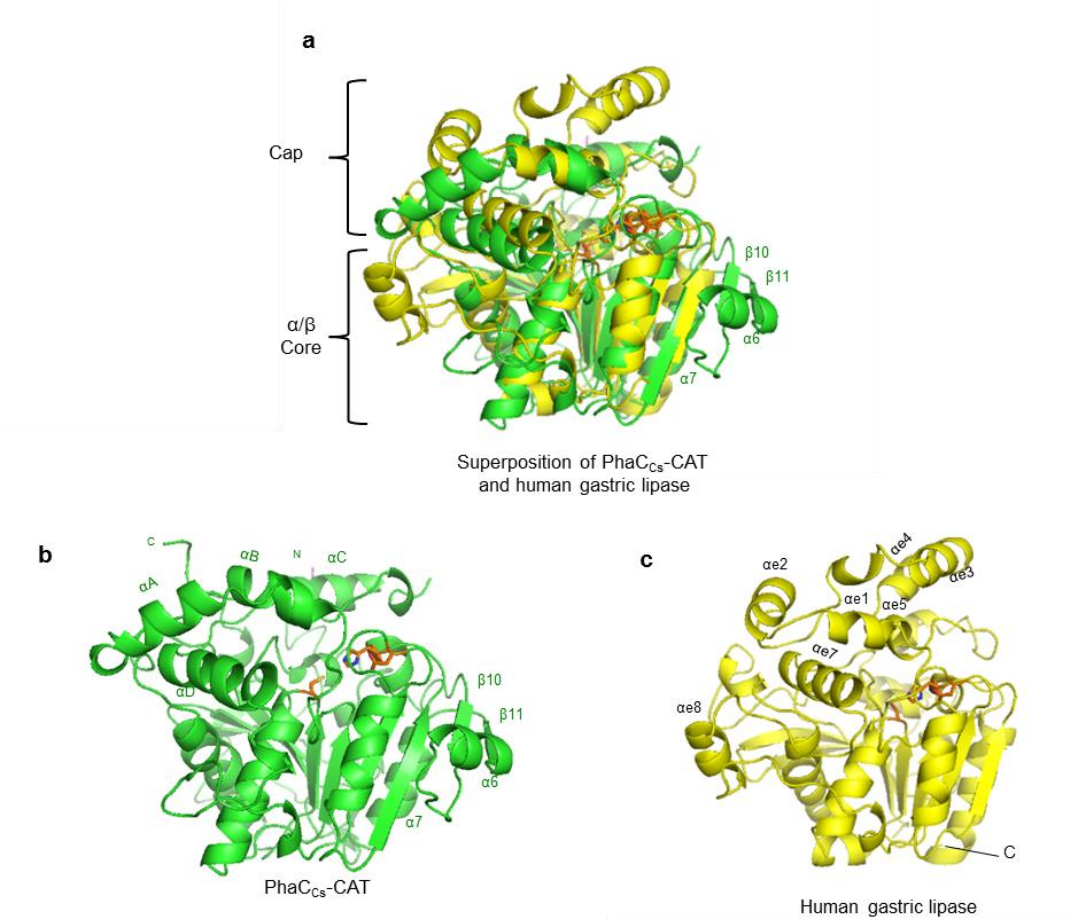


Figure 8. Structural comparison of the catalytic domains of PhaC$_{Cs}$ and lipase.

(**a**) PhaC$_{Cs}$-CAT overlaid on human gastric lipase. The α/β core subdomains are well overlapped with a small r.m.s. deviation (1.02 Å), whereas the CAP subdomains are poorly overlapped. The α/β core subdomain of PhaC$_{Cs}$-CAT contains an additional segment comprising β10-α6-β11-α7.

(**b**) As in **a**, but for a side-view of PhaC$_{Cs}$-CAT determined in the current study.

(**c**) As in **b**, but for human gastric lipase (PDB code 1HLG).

3.1.1 Catalytic triad of PhaC$_{Cs}$-CAT.

PhaC$_{Cs}$ possesses a catalytic triad, comprised of Cys291, Asp447, and His477, which is important for its catalytic activity and conserved in all members of PHA synthases. The catalytic Cys291 residue is located at the turn between β6 strand and α3 helix. This turn contains a conserved lipase box-like motif Gx$_1$Cx$_2$G (where x$_1$ is Phe290 and x$_2$ is Val292 in PhaC$_{Cs}$) and forms a "nucleophilic elbow" found in α/β hydrolases, holding the catalytic cysteine residue at the top position of the elbow (**Fig. 9a**). In the current structure, Cys291 forms the catalytic triad with His477 (located at the N-end of α5 helix) and Asp447 (located at the loop between β8 strand and α4 helix, β8-α4 loop). Previous studies have shown the importance of the catalytic triad as its mutation diminished catalytic activity[67]. The residue next to the catalytic cysteine, Val292, is found to form the oxyanion hole with its main-chain amide group. The catalytic triad is located at the N-terminal edge of the core parallel β-sheet. Together, five α-helices (α1-α5) point the N-terminal edge with their N-terminal ends. This phenomenon might contribute to the stabilization of the negatively charged molecule such as acyl-CoA. However, in the current structure, the active site is covered by the CAP subdomain, blocking the substrate 3HB-CoA access to the active site (**Fig. 6d**).

The other member of the catalytic triad, Asp447, is stabilized by three hydrogen bonding formed with the main chain NH groups of β8-α4 loop residues His448, Ile449 and Val450 (3.1 Å, 2.8 Å and 2.9 Å, respectively) (**Fig. 9a**). At the same time, the Asp447 side chain forms another hydrogen bond with His477 (Nδ) at a distance of 2.7 Å. In the catalytic pocket, a water molecule was observed forming a hydrogen bond with the catalytic His477 with the ring nitrogen atom (Nε). The same Nε atom is

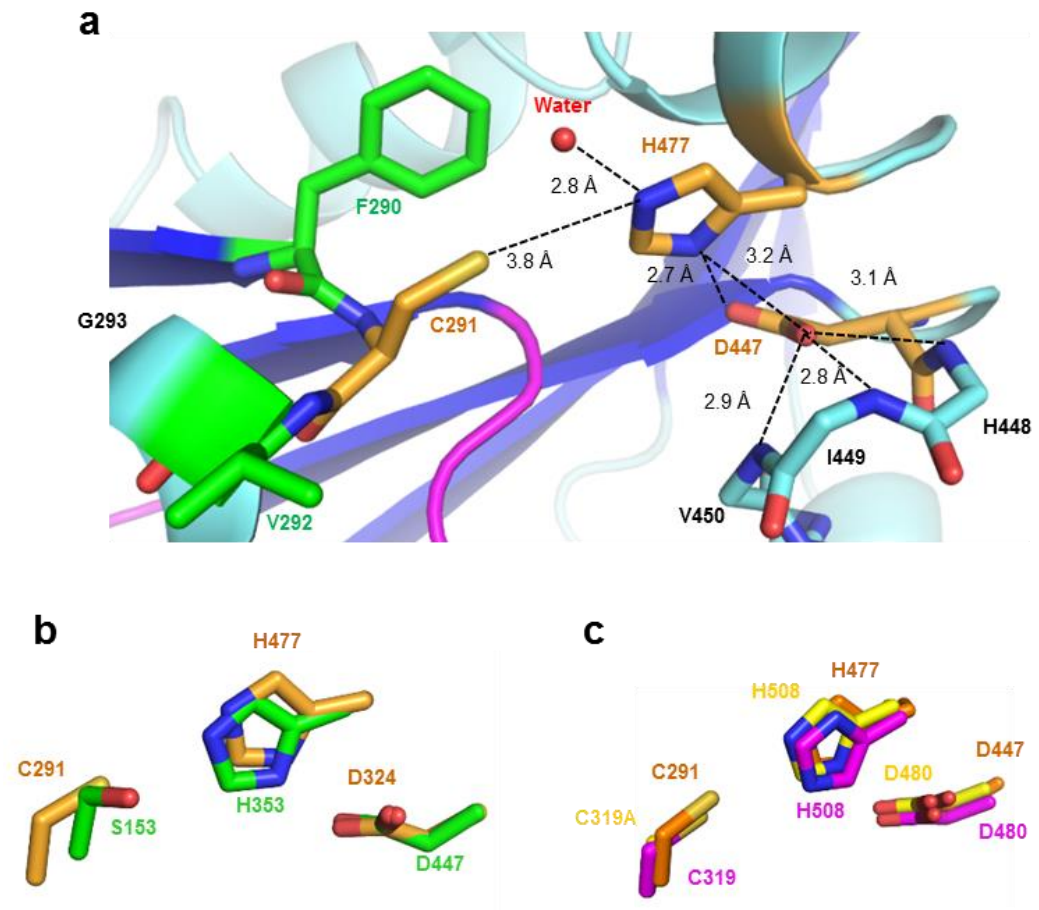


Figure 9. Catalytic triad of PhaC$_{Cs}$-CAT.

(**a**) A close-up view of the catalytic site of PhaC$_{Cs}$-CAT. The catalytic triad residues (orange) comprise Cys291, His477 and Asp447, with a hydrogen bond between His477 and Asp447. The catalytic center Cys291 is located at the nucleophilic elbow sandwiched with nonpolar residues Phe290 and Val292 (green). Hydrogen bonds are indicated by broken lines.
(**b**) Overlay of the catalytic triad of PhaC$_{Cs}$-CAT (orange) onto human gastric lipase (1HLG in green).
(**c**) Overlay of the active site with the catalytic triad of PhaC$_{Cs}$-CAT (orange) onto those of PhaC$_{Cn}$-CAT (5T6O in yellow; 5HZ2 in magenta). The imidazole ring of the His residue of PhaC$_{Cn}$-CAT (magenta) is flipped from the other two.

also position at a close distance (3.8 Å) to the sulfur atom (S_γ) of Cys291.

The geometry of the residues of the catalytic triad well corresponds to that of lipases (**Fig. 9b**), suggesting a similar catalytic mechanism to these residues catalyzing the esterification reaction. In the biosynthesis of PHA catalyzed by PhaC, active center Cys291 involved in the nucleophilic attack of the thioester of acyl-CoA such as 3HB-CoA. The attack should be accelerated by His477 through the deprotonated thiol group of Cys291. In the current structure, Asp447 assists His477, by pulling the proton from the His477 through forming the hydrogen bonding, enhancing the deprotonation of Cys291. Also, Asp447 may act as a general base catalyst by accelerating deprotonation of the hydroxyl group of the substrate acyl-CoA[38,68].

As described below, in the open form CoA bound PhaC_{Cs}-CAT structure, the Cys291 and Asn447 (Asp in wild-type) were similar to the free form structure. The major difference observed at His477 where it had flipped out from the active site. This phenomenon is in accordance to the previous proposed mechanism where the His477 is responsible for the Cys291 activation and reaction initiation.

3.1.2 Clusters of water molecules at the active site.

The presence of water molecules inside and outside of the PhaC_{Cs}-CAT were clearly displayed in this high-resolution structure (1.48 Å). Several clusters of water molecules were observed in the cavity between the CAP and α/β core subdomains. This observation suggests the possibility of the structural movement of CAP subdomain in solution. PhaC_{Cs} may undergo a dynamic structural change during catalytic reaction. Closer analysis at the cavity around active center Cys291 revealed

water molecules covered by part (Pro327-Pro386) of the CAP subdomain (**Fig. 10a**). This part of the CAP subdomain contains ηA, αA, ηB, αB' and ηB' helices and corresponds to the segment Thr355-Pro419 of PhaC$_{Cn}$. The water molecules observed could be divided into two groups (Site A and Site B) separated by the nucleophilic elbow (**Fig. 10b**). Site A formed by Val292 from the nucleophilic elbow and Val295 from α3 helix with two polar residues, Tyr412 from αD helix and His324 from β7-ηA loop, and other nonpolar residues (Leu321 from β7-ηA loop, Ile449 and Val450 from β8-α4 loop, Met407 from αD helix and Ile219 from β4-η1 loop). The other nonpolar Site B formed by Phe290 from the nucleophilic elbow with other nonpolar residues (Pro216, Pro217, Ile219 from β4-η1 loop, Leu224 and Met225 from η1 helix, and Trp392 from αC helix). Asn220 is the only polar residue that faces this water cluster.

3.1.3 Conformational changes in the CAP subdomain.

The amino acid sequences of PhaC$_{Cs}$ shared a sequence identity at 46% with PhaC from *Cupriavidus necator* (PhaC$_{Cn}$), one of the most studied class I PHA synthases. The structural comparison of current PhaC$_{Cs}$-CAT with the recently reported structure of PhaC$_{Cn}$-CAT[69,70] may provide a better understanding of this enzyme. As expected, the α/β core subdomains resemble each other and the relative positions of the Cys-His-Asp residues of the catalytic triads are well conserved, whereas the imidazole ring of the His residue is flipped in the structure determined by the KNU group[70] (**Fig. 9c**). Interestingly, unexpected structural deviations in the CAP subdomain are observed in the structural comparison (**Fig. 11**). Part of the CAP subdomain, LID region (residues Pro327–Pro386 in magenta) of the PhaC$_{Cs}$-CAT, showed obvious conformational deviations. The LID region of PhaC$_{Cs}$-CAT corresponds to the

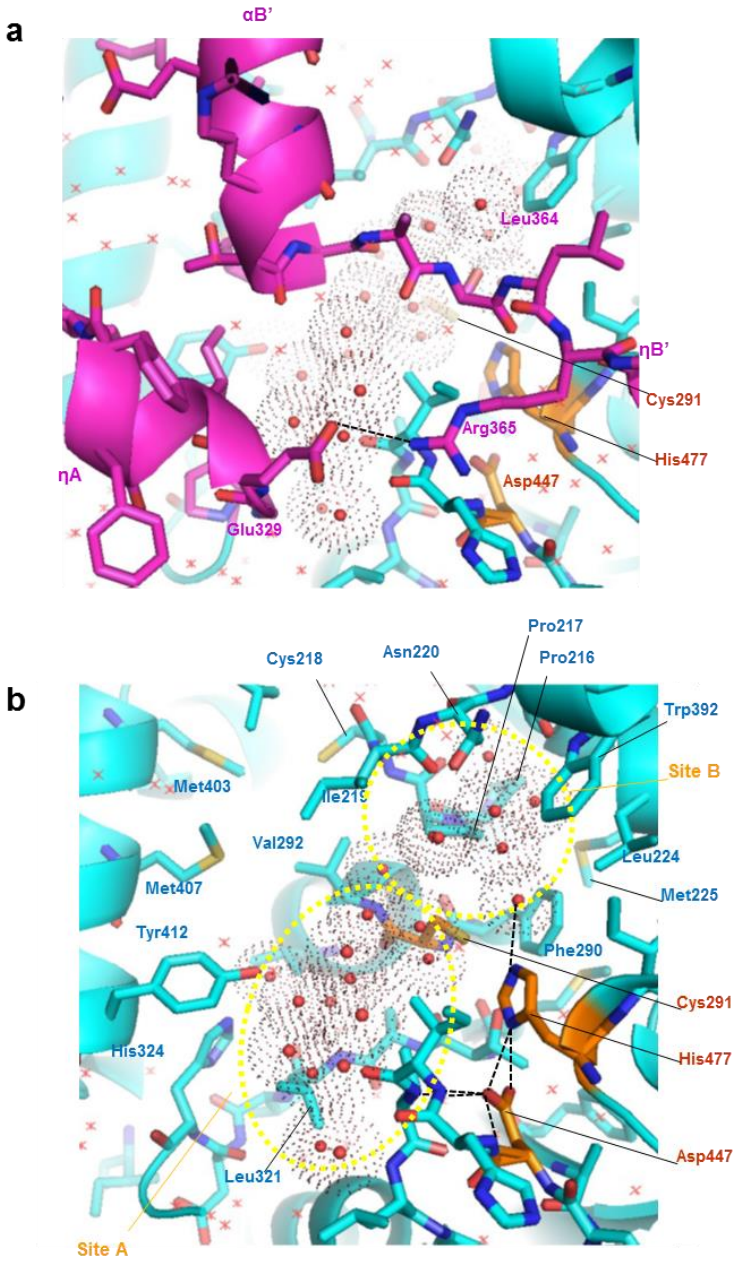


Figure 10. Water molecules at both sides of the nucleophilic elbow of PhaC$_{Cs}$-CAT.

(**a**) A cluster of water molecules (red balls with dotted surfaces) is located at the active site cavity around the nucleophilic elbow. The cavity is covered by the LID region of the CAP subdomain (magenta). Other water molecules are shown as red crosses. Color codes are the same as in Figure 5. The hydrogen bond between Glu329 and Arg365 is indicated by a broken line.

(**b**) As in **a**, but without the LID region. Hydrogen bonds involving the catalytic residues are shown as broken lines. Water molecules are divided into two groups, one group at Site A and the other at Site B. The cavity is mostly hydrophobic although Tyr412 and His324 are located at Site A and Asn220 at Site B. Hydrogen bonds are indicated by broken lines.

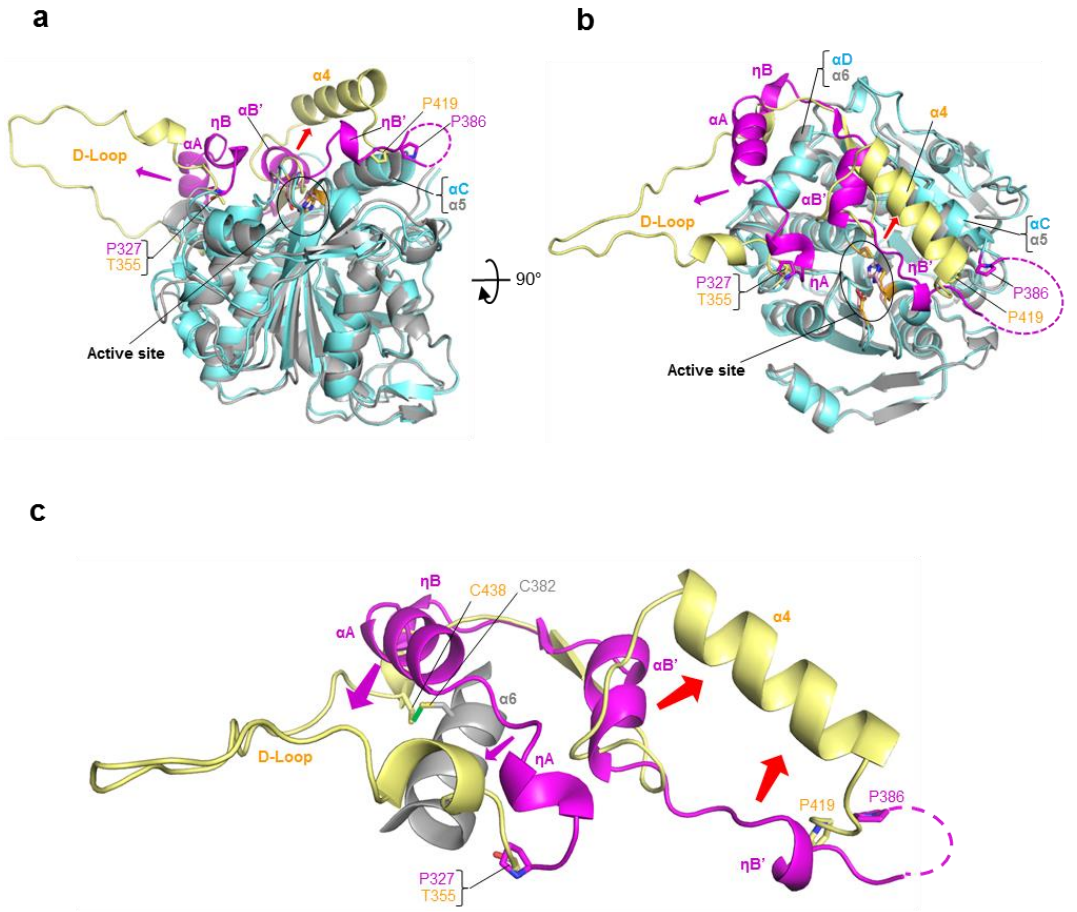


Figure 11. Conformational changes in the CAP subdomain as revealed by structural comparison of PhaC$_{Cs}$-CAT and PhaC$_{Cn}$-CAT structures.

(**a**) A side-view of the overlay between PhaC$_{Cs}$-CAT (cyan and magenta) and PhaC$_{Cn}$-CAT (gray and yellow) with an overall r.m.s. deviation of 2.7 Å for Cα carbon atoms. The major conformational deviations are observed in part of the CAP subdomain, the LID region (residues Pro327–Pro386 in magenta) of PhaC$_{Cs}$-CAT, which corresponds to the segment Thr355–Pro419 (yellow) of PhaC$_{Cn}$-CAT. Arrows indicate conformational transitions of the LID region from PhaC$_{Cs}$-CAT to PhaC$_{Cn}$-CAT.

(**b**) As in **a**, but a top-view down to the active site. A narrow path to the active site was found in PhaC$_{Cn}$-CAT, whereas the path is covered by the LID region in PhaC$_{Cs}$-CAT.

(**c**) To clarify the conformational transitions found by the structural comparison of PhaC$_{Cs}$-CAT and PhaC$_{Cn}$-CAT, overlay of the LID regions of the two structures are shown. Unfolding of ηA, αA and ηB helices of PhaC$_{Cs}$-CAT into a long flexible D-loop found in PhaC$_{Cn}$-CAT is indicated by magenta arrows, and refolding of αB' and ηB' helices and their linker loop of PhaC$_{Cs}$-CAT into a long α4 helix in PhaC$_{Cn}$-CAT by red arrows. The disulfide bond between Cys328 and Cys438 of PhaC$_{Cn}$-CAT is shown in green.

segment Thr335-Pro419 of PhaC$_{Cn}$-CAT. These conformational differences could be explained by two dynamic conformational transitions (**Fig. 11c**). One event involved the unfolding of ηA, αA and ηB helices of PhaC$_{Cs}$-CAT into a long flexible D-loop found in PhaC$_{Cn}$-CAT. The other involves a folding of αB' and ηB' helices and their linker loop of PhaC$_{Cs}$-CAT into the long α4 helix in PhaC$_{Cn}$-CAT with a large shift in positions that uncovers a path to the active site. Interestingly, the segment (Leu402–Asn415) forming α4 helix of PhaC$_{Cn}$-CAT has a conserved sequence among Class I and II PHA synthases (**Fig. 7**, green box), whereas the corresponding segment (Leu369–Lys382) of PhaC$_{Cs}$-CAT displays a disordered structure. This difference could be caused in part by the difference in dimerization modes as discussed below.

3.1.4 An artificial disulfide bond formation may induce the conformational change in the CAP subdomain of PhaC$_{Cn}$-CAT.

In the structural comparison of PhaC$_{Cn}$-CAT, one disulfide bond is found between Cys382 and Cys438[69,70] (**Fig. 11c**). However, these Cys residues are not universally conserved and such disulfide bond is not found in PhaC$_{Cs}$-CAT (**Fig. 7**). In the PhaC$_{Cn}$-CAT structures, Cys382 is in the insertion segment specific to PhaC$_{Cn}$ and located in part of the D loop of the LID region, which displays a distinct conformation from our PhaC$_{Cs}$-CAT structure. While Cys438 is located at a common helix corresponding to αD helix of our PhaC$_{Cs}$-CAT structure. It is likely that disulfide bond formation induces the long D-loop and stabilizes the partially open conformation of the CAP subdomain, which is distinct from the closed conformation of our PhaC$_{Cs}$-CAT structure. It should be noted that PhaCs are cytoplasmic enzymes that act under reduced conditions of the redox potential, indicating that artificial

formation of the disulfide bond may have occurred during crystallization. However, it may be possible that artificial disulfide bond formation contributes in trapping the partially open state of the CAP subdomain in an open-closed equilibrium in solution.

3.1.5 Dimeric structure of PhaC$_{Cs}$-CAT.

PHA synthases are known to exist in an equilibrium between monomer and dimer in solution and are believed to possess full catalytic activity in the dimeric form[35,71,72]. In an effort to analyze the equilibrium of PhaC$_{Cs}$ in the absence or presence of substrate (DL-3HB-CoA), analytical ultracentrifugation (AUC) was employed utilizing a sedimentation velocity method (**Fig. 12a**). The analysis revealed that at low concentration (5 µM), free PhaC$_{Cs}$ (full-length, 63.4 kDa) exists in an equilibrium between monomeric and dimeric forms, with the monomeric form being dominant at this concentration. In the presence of substrate at a 10-fold molar ratio (50 µM), the equilibrium is shifted to include a higher order oligomer (presumably a tetramer). The monomer-dimer equilibrium was further analyzed by size exclusion chromatography (SEC) at a concentration of 30 µM with full-length PhaC$_{Cs}$ and PhaC$_{Cs}$-CAT (**Fig. 12b**). In the absence or presence of substrate (500 µM), full-length PhaC$_{Cs}$ exists in a monomeric-dimeric equilibrium with the monomeric form being dominant, which is similar to the results from analytical ultracentrifugation. Compared with this result, the monomeric form of PhaC$_{Cs}$-CAT is more dominant and only a trace amount of the dimeric form was detected in SEC. Further analyses of the monomeric-dimeric equilibrium by AUC showed that the dimeric form of

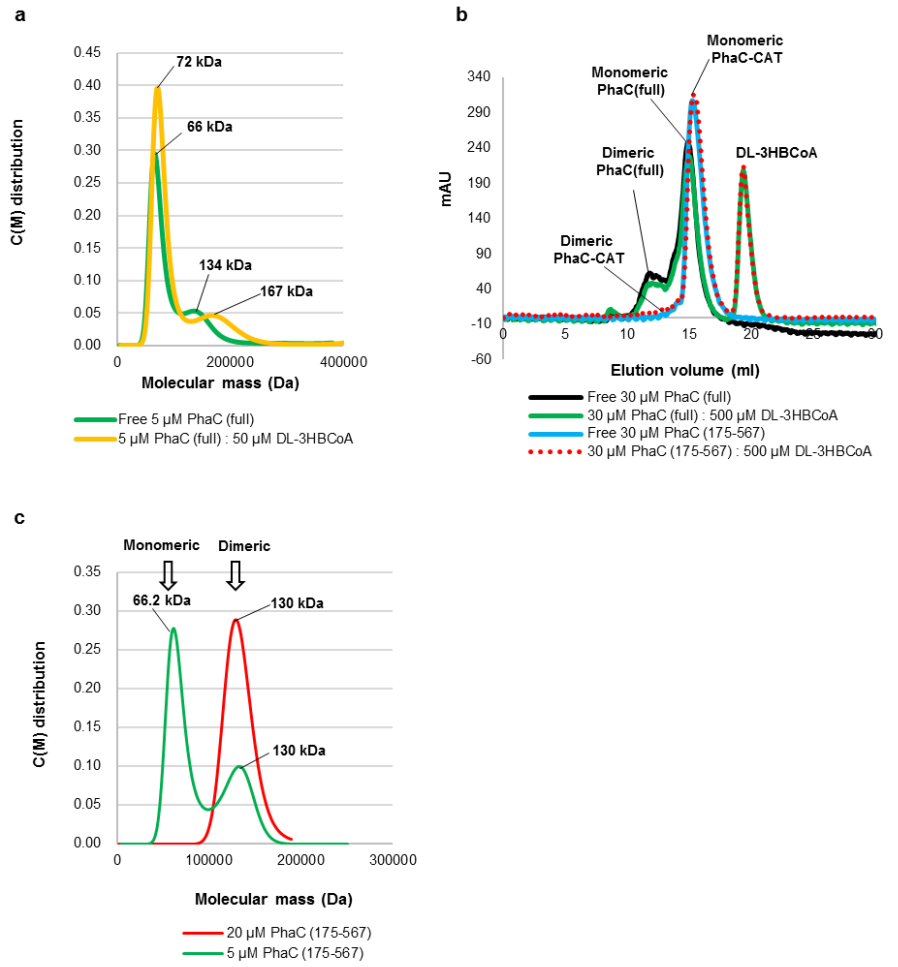


Figure 12. Oligomerization equilibrium of PhaC in solution.

(**a**) Analytical ultracentrifugation (AUC) analyses of PhaC$_{Cs}$ (full-length, 63.4 kDa) using a sedimentation velocity method. The equilibria between monomeric and dimeric forms in solution were observed in the absence or presence of substrate (DL-3HB-CoA). The monomeric form is dominant at 5 µM PhaC$_{Cs}$ in 10 mM Tris-HCl (pH 8.0), 100 mM NaCl and 3 mM β-ME at 20˚C, with a small portion being present in the dimeric form. In the presence of substrate, a higher (presumably tetrameric) form appears.

(**b**) Size exclusion chromatography (SEC) of PhaC$_{Cs}$ (full-length) and PhaC$_{Cs}$(175-567), which is PhaC$_{Cs}$-CAT, in the absence or presence of substrate DL-3HB-CoA. The elution profiles of PhaC$_{Cs}$ and PhaC$_{Cs}$-CAT were compared.

PhaC$_{Cs}$ (full-length) exists in monomeric and dimeric (dominant) forms. Addition of DL-3HB-CoA induced only a small change with the appearance of a faint peak, presumably that of the tetrameric form. In contrast to these equilibria, PhaC$_{Cs}$-CAT exists in a monomeric form at a concentration of 30 µM both in the absence or presence of substrate.

(**c**) AUC analyses of PhaC$_{Cs}$-CAT using a sedimentation velocity method. The dimeric form of PhaC$_{Cs}$-CAT was detected at 5 µM (green) and significantly stabilized at a high concentration (20 µM, red).

PhaC$_{Cs}$-CAT exists and is stabilized at a high protein concentration (20 μM) (**Fig. 12c**). In SEC, monomers are separated from each other in the resin and may fail to re-associated to form dimers again. Thus, the monomeric form seems to be apparently stabilized in SEC. The asymmetric unit contains two PhaC$_{Cs}$-CAT molecules (mol A and mol B), which form a face-to-face dimer with a pseudo dyad axis (**Fig. 13a**). It is likely that the high protein concentrations (~0.3 mM) in our crystallization reproduce the catalytically relevant form in the crystal. It is noteworthy that in the current dimer, the N-terminal ends of both protomers are exposed on the same side of the dimer, implying that the N-terminal subdomain, which is absent in the current PhaC$_{Cs}$-CAT, is important for dimer stabilization and that dimerization could be mediated by direct contacts of N-terminal domain from both protomers (**Fig. 13b**).

In the PhaC$_{Cs}$-CAT dimer, the CAP subdomains mediate the intermolecular contacts at the dimer interface, although part of the core domains also participate in formation of the interface. The intimate contact areas comprise ηA (Glu329, Phe332, Phe333) and ηB'(Arg365) helices from each CAP subdomain and the β8-α4 loop (His448, Leu451) from each core subdomain. One of the contact areas (Contact area I) forms a hydrophobic cluster with nonpolar residues from ηB'-αC loop (Leu369, Trp371) and αC helix (Pro386, Phe387, Leu390) from both protomers. The segment (Tyr373-Thr383) between ηB'-αC loop (Leu369, Trp371) and αC helix are disordered in our crystal (dashed lines) (**Fig. 13c**). Contact area II contains salt bridges (between Arg365 and Glu329) buried inside the interface formed by nonpolar residues of ηA helix (Phe332, Phe333) and β8-α4 loop (His448, Leu451) from both protomers (**Fig. 13d**). Since β8-α4 loop contains one of the catalytically active

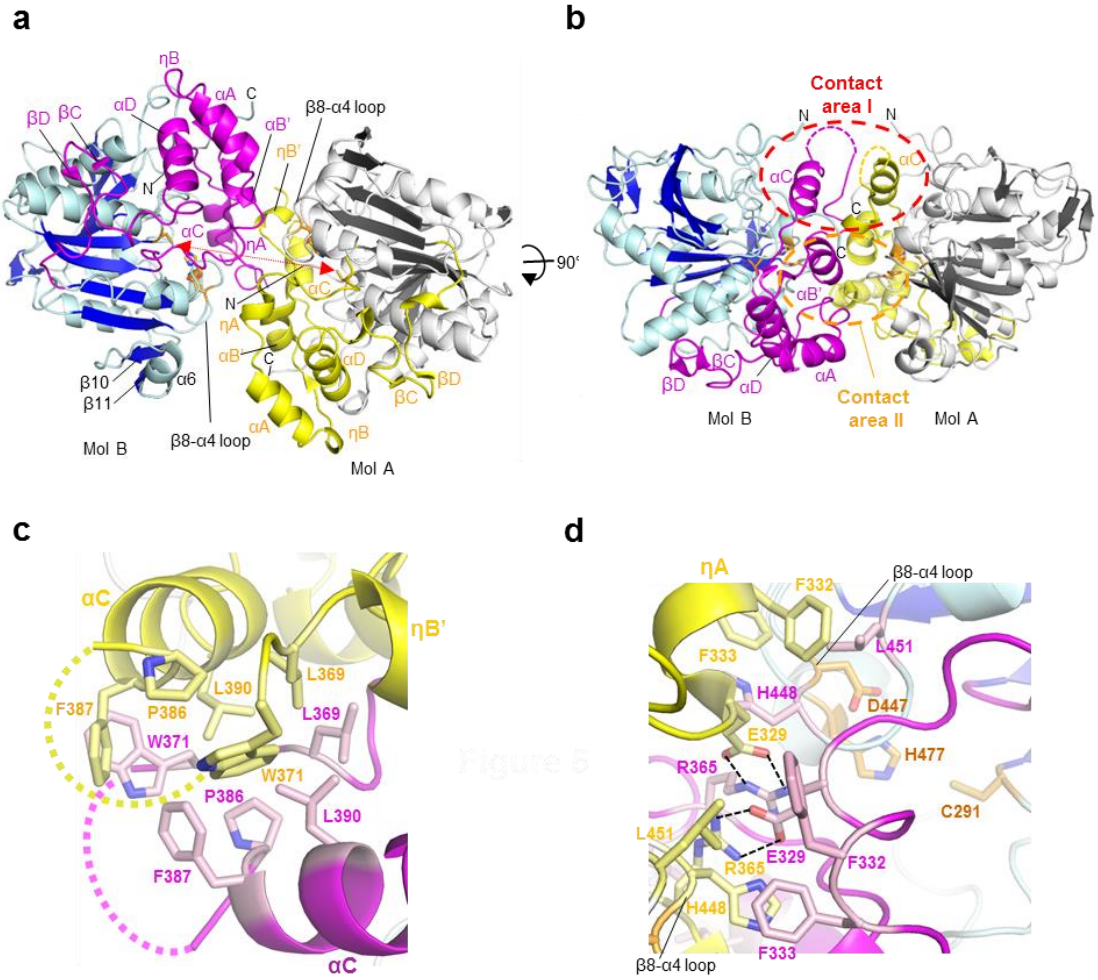


Figure 13. PhaC$_{Cs}$-CAT forms a dimer in the crystal structure.

(**a**) A top-view of the PhaC$_{Cs}$-CAT dimer along the pseudo-dyad axis. The color codes of mol B are the same as in Fig. 1, with mol A core (gray) and CAP (yellow) subdomains shown. The distance between the catalytic cysteine residues is 28.1 Å (red double-headed arrow).

(**b**) As in **a**, but the side-view. Contact areas I (red) and II (orange) are indicated by broken circles.

(**c**) Contact area I forms a hydrophobic cluster with nonpolar residues from ηB'-αC loop (Leu369, Trp371) and αC helix (Pro386, Phe387, Phe390) from both protomers. The segment (Asn372-Thr383) between ηB'-αC loop (Leu369, Trp371) and αC helix are disordered in our crystal (dashed lines).

(**d**) Contact area II contains salt bridges (Arg365---Glu329) buried inside the interface formed by nonpolar residues of ηA helix (Phe332, Phe333) and β8-α4 loop (His448, Leu451) from both protomers. Hydrogen bonds are indicated by broken lines.

residues, Asp447, dimer formation could have some effects on enzymatic activity. The catalytic cysteine residues in the PhaC$_{Cs}$-CAT dimer are distant (28.1 Å for the distance between the S_γ atoms of Cys residues), and no clear path between them was found in the current structure.

3.1.6 Conformational changes in the CAP subdomain induce rearrangement of the dimer interface.

Since the CAP subdomain is a major part of the dimer interface, rearrangement of dimer association by conformational changes in the CAP subdomain is evident when comparing the PhaC$_{Cs}$-CAT and PhaC$_{Cn}$-CAT dimeric structures (**Fig. 14**). If one protomer of each dimer is superimposed, the other protomer is swung by ~40 Å with a rotation of ~120°. The distance between the catalytic Cys residues in the PhaC$_{Cn}$-CAT dimer is 33.3 Å, which is longer than the distance (28.1 Å) in the PhaC$_{Cs}$-CAT dimer. The distance between the N-termini of the protomers in the PhaC$_{Cn}$-CAT dimer is longer (55.1 Å) than that (19.2 Å) of the PhaC$_{Cs}$-CAT dimer, although the N-termini are still located on the same side of the PhaC$_{Cn}$-CAT dimer surface (**Fig. 14b, c**). This implies that reorganization of the dimer interface induce`d by refolding/unfolding of the LID region of the CAP subdomain may be facilitated without dissociation of the protomers.

3.2 Overall structure of open form CoA-bound PhaC$_{Cs}$-CAT complex.

It is generally depicted that Coenzyme A (CoA) has an inhibitory effect on the PhaC catalytic action as a by-product released from the active site after the acyl moiety was cleaved from the substrate acyl-CoA[73,74]. Two independent studies have

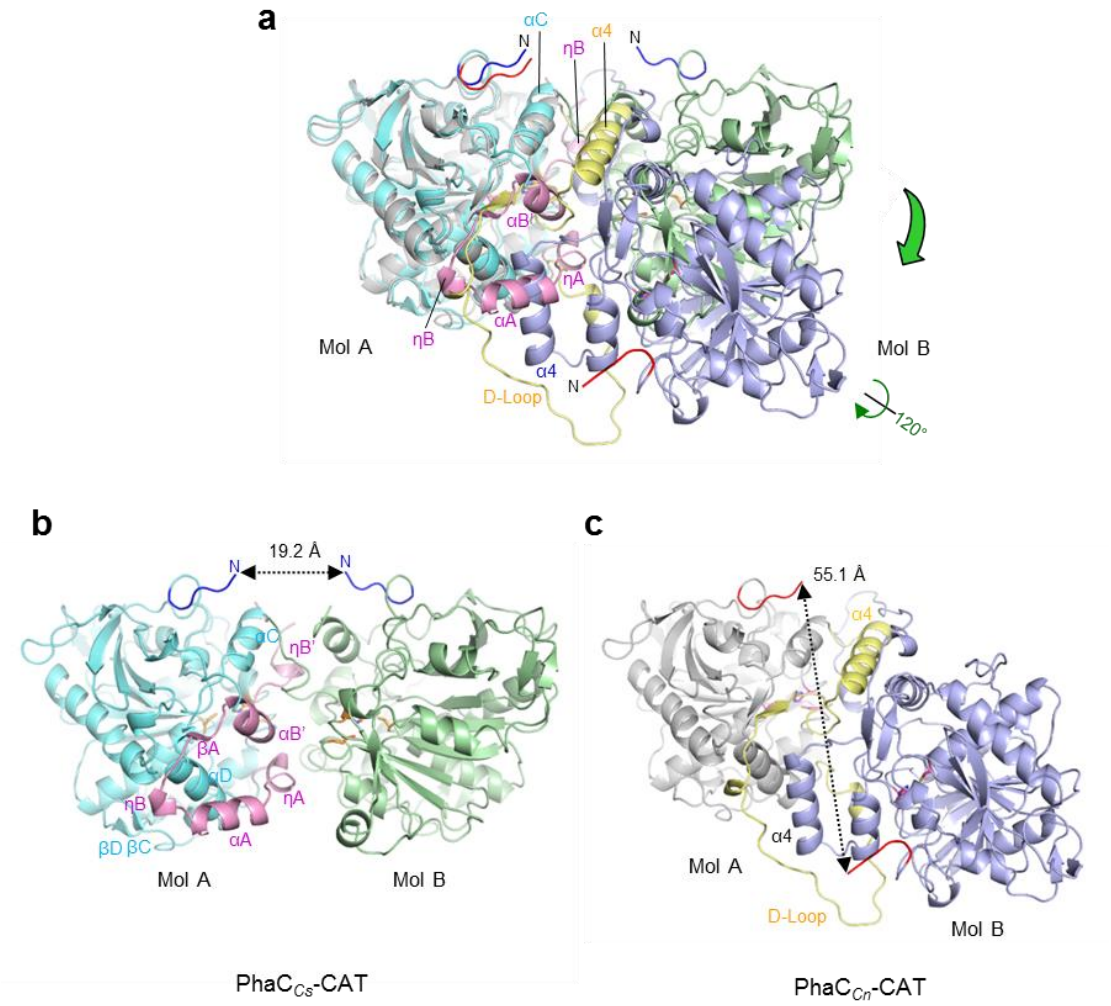


Figure 14. Structural comparison of PhaC$_{Cs}$-CAT and PhaC$_{Cn}$-CAT dimers.

(**a**) Overlay of the PhaC$_{Cs}$-CAT dimer (cyan and green) on the PhaC$_{Cn}$-CAT dimer (gray and purple). One protomer (cyan) of the PhaC$_{Cs}$-CAT dimer is superimposed on one protomer (gray) of the PhaC$_{Cn}$-CAT dimer. The distance between the catalytic cysteine residues in the PhaC$_{Cs}$-CAT dimer is 28.1 Å, while that in the PhaC$_{Cn}$-CAT dimer is 33.3 Å.

(**b**) The PhaC$_{Cs}$-CAT dimer. The distance between the N-termini is 19.2 Å.

(**c**) The PhaC$_{Cn}$-CAT dimer with mol A in the same orientation as mol A in the PhaC$_{Cs}$-CAT dimer as in **b**. The dimer interface is reorganized by refolding/unfolding of the LID region with accompanying rotation of one protomer. The distance between the N-termini is 55.1 Å.

showed the importance of the CoA moiety for PHA synthase activity as class I PhaC$_{Cn}$ could not catalyze substrate analogs that lack of the adenosine 3′, 5′-bisphosphate or 3′ phosphate moiety in 3HB-CoA[73,74]. Another inhibition study using class III PhaC-PhaE complex from *Allochromatium vinosum* (PhaCE$_{Av}$) shows a similar result that the CoA moiety might be essential for substrate binding[75]. In our pursuit of an intermediate complex, co-crystallization with CoA using various catalytic mutants were attempted. Among the trials, PhaC$_{Cs}$-CAT of Asp447 to Asn mutation able to form a complex crystal with CoA (hereafter referred to as PhaC$_{Cs}$-CAT+CoA). The PhaC$_{Cs}$-CAT+CoA complex structure (3.10 Å) was solved by molecular replacement using free form PhaC$_{Cs}$-CAT structure as a search model (**Table 4).**

The crystal structure of PhaC$_{Cs}$-CAT+CoA complex revealed that the enzyme forms a dimer comprised of a fully closed form protomer (mol A) and a fully open form CoA-bound protomer (mol B) by interacting each other with their CAP subdomain (**Fig. 15a**). This observation is in agreement with my previous suggestion that conformational changes in PhaC from closed to open may happen without dissociation of its dimeric conformation.

At present, four conformations of PhaC-CAT structures were determined, namely fully closed (PhaC$_{Cs}$-CAT+CoA, mol A) (**Fig. 15b**), semi-closed (the free form PhaC$_{Cs}$-CAT) (**Fig. 15c**), partially open (the free form PhaC$_{Cn}$-CAT) (**Fig. 15d**), and fully open (PhaC$_{Cs}$-CAT+CoA, mol B) forms (**Fig. 15e**). The major difference between the fully- and semi-closed forms is the presence of a long αB helix in the fully closed form (**Fig. 15b**). While in the semi-closed form, part of the αB helix (373-YVVNNYLLGKT-383) was disordered (**Fig. 15c**). This long αB helix of the

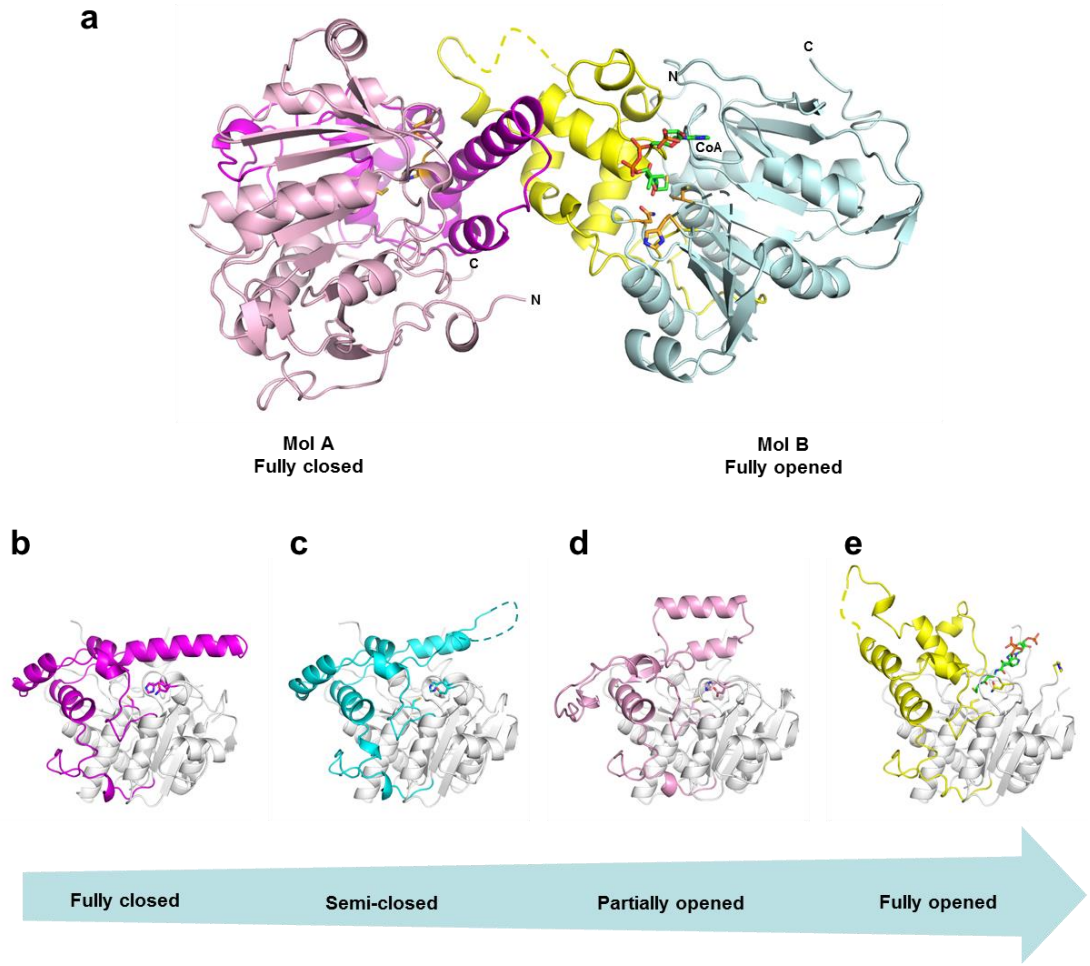


Figure 15. Complex structure of CoA-bound PhaC$_{Cs}$-CAT and various conformations.

(**a**) Complex structure contain closed (pink and magenta) and CoA-bound open (white and yellow) forms of PhaC$_{Cs}$-CAT. The dimer inteface is comprised of CAP subdomains (magenta and yellow) from each protomers. The Coenzyme A is shown as green sticks.

(**b**)-(**e**) All available PhaC-CAT structures in different conformations. The CAP subdomain was coded in various colors: fully closed (PhaC$_{Cs}$-CAT+CoA mol A), magenta; semi-closed (PhaC$_{Cs}$-CAT+CoA mol A), cyan; partially open (PhaC$_{Cn}$-CAT mol B), pink; fully open (PhaC$_{Cs}$-CAT mol B), yellow.

fully closed form (mol A, **Fig. 16. a, b**) interacts with part of the CAP subdomain from the fully open form (mol B, **Fig. 16c, d**) through hydrophobic and salt-bridging interactions (**Fig. 16e**). In the open form, three segments (residues 448-488, 548-567 in the core subdomain and 349-355 in the CAP subdomain) were failed to be determined because of poor electron density. While the core subdomain largely remained similar as the fully-/semi-closed form, the catalytic His477 is shifted out from the catalytic site. The LID region (Pro327-Pro386, as shown in **Fig. 7**) of the fully open form experienced conformational changes and allowed CoA entry into the active site (**Fig. 16c, d**).

3.2.1 CoA-binding tunnel.

A clear electron density is observed in the active site of the open form mol B of the PhaC$_{Cs}$-CAT+CoA structure, which fits the entire CoA molecule (**Fig. 17a**). To avoid model bias, omit electron density map (F_o-F_c) of CoA was calculated at contour level of 3σ (**Fig. 17b)**. The omit map well fits onto the CoA molecule, whereas no such electron density was found in the active site of the fully closed form of the PhaC$_{Cs}$-CAT+CoA complex.

The extended pantetheine arm of CoA interacts with PhaC$_{Cs}$-CAT through hydrogen bonds in the active site tunnel (**Fig. 17c, d)**. The catalytic center Cys291 (S_γ) holds the end of the β-mercaptoethylamine tail (S_1) of CoA through hydrogen bonding (3.1 Å) (**Fig. 17c, d**). The mutated Asn447 interacts with CoA (O_5) *via* two hydrogen bonds with $O_{\delta1}$ (2.9 Å) and $N_{\delta2}$ (2.7 Å) (**Fig. 17c, d**). The main-chain O of Leu380 and Gly381 (αB-αC loop) interact with CoA (N_4) [3.4 Å] and CoA (N_8) [3.0 Å], respectively (**Fig. 17c, d**). The wall of CoA-binding tunnel also formed by

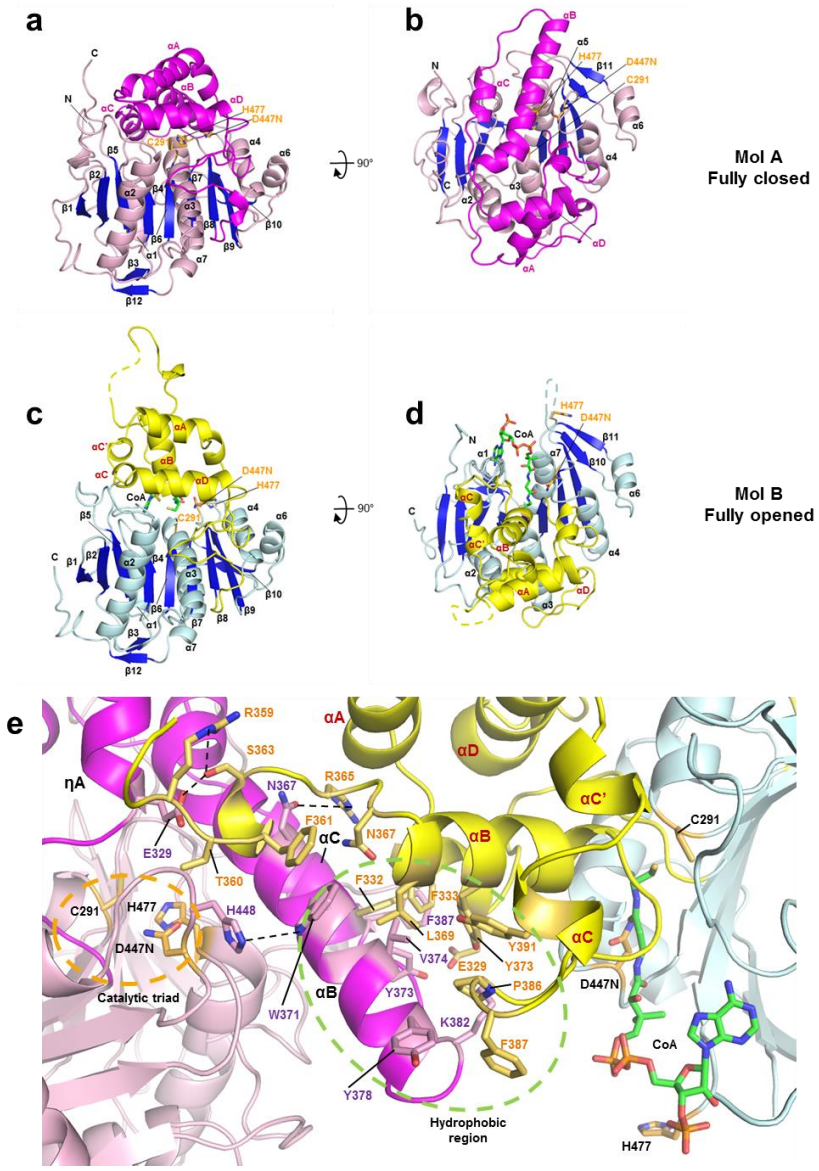


Figure 16. Two forms found in the complex structure of CoA-bound PhaC*Cs*-CAT and the dimer interface.

(**a**) A side view of mol A from PhaC*Cs*-CAT+CoA, representing a fully closed form where the disordered region observed in the free form PhaC*Cs*-CAT is folded into a long αB helix.

(**b**) As in **a**, but a top view. The catalytic site is completely covered by the CAP subdomain, particularly by αB and αC helices.

(**c**) A side view of mol B from PhaC*Cs*-CAT+CoA, representing a fully open form. Dynamic conformational transition of CAP subdomain (yellow) allows CoA to enter into the active site.

(**d**) As in **c**, but a top view. CoA (green) is located at the catalytic site and His477 is shifted out from the catalytic site.

(**e**) Dimer interface of PhaC*Cs*-CAT+CoA. At the direct contact region of the CAP subdomains from the closed (magenta) and open (yellow) form PhaC*Cs*-CAT, the side-chains of residues involved in the contacts and CoA are shown in sticks model.

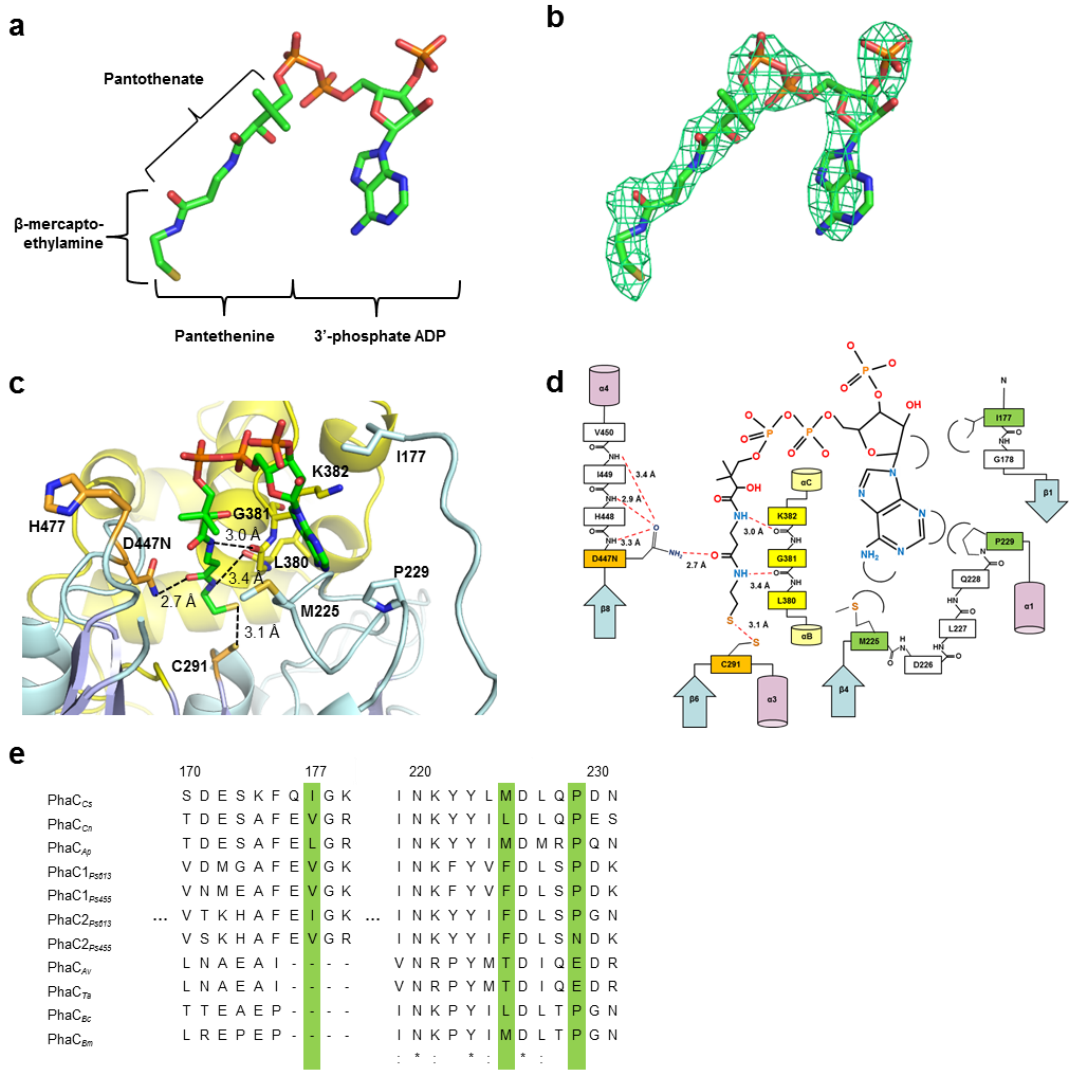


Figure 17. Coenzyme A (CoA) interaction with PhaC$_{Cs}$-CAT.

(**a**) The CoA consists of pantetheine arm and a 3'-phosphate ADP moiety.

(**b**) The model of CoA in the omit electron density map (F_o-F_c) at a contour level of 3σ.

(**c**) Open form PhaC$_{Cs}$-CAT (core subdomain, cyan; CAP subdomain, yellow) interacts with CoA (sticks, green). Catalytic C291 and D447N interact with CoA through hydrogen bonds. Main-chain oxygen atoms from L380 and G381 form hydrogen bonds with N_4 and N_8 of CoA.

(**d**) Schematic representation of the interactions between PhaC$_{Cs}$-CAT and CoA. Catalytic residues are coded in orange; Residues from the CAP subdomain are color-coded in yellow; Hydrophobic residues binding adenine group are coded in green.

(**e**) Sequences alignment revealed the conservation of hydrophobic residues correspond to I177, M225, and P229 from PhaC$_{Cs}$.

residues from both the core (side-chain of Leu224, Met225, Ile449, Val450) and CAP subdomain (main-chain of Leu379, Leu380, Gly381) (**Fig. 17c, d**). The adenine moiety of CoA was fit into a hydrophobic cavity derived from Ile177, Met225, and Pro229 (**Fig. 17c, d**). The corresponding residues are highly conserved in hydrophobicity among class I, II PHA synthases (**Fig. 17e**). Previous studies indicate the importance of CoA moiety in substrate binding for the PhaC catalytic activity[73-75]. The fact that CoA lacking the 3′ phosphate and 3′-P-5′-P-ADP showed no inhibitory effect on the PhaC activity[73] suggests both the 3′ phosphate and adenosine 3′, 5′-bisphosphates moiety of CoA are required for CoA binding. The chemically favorable interactions between CoA and PhaC$_{Cs}$-CAT suggest the current complex structure of PhaC$_{Cs}$-CAT+CoA is a biological relevant one.

3.2.2 Comparison of active site of the open and closed PhaC$_{Cs}$-CAT.

PhaC possess a conserved catalytic triad comprised of Cys, Asp, and His residues, which are important for its catalytic activity. Since the catalytic triad is the core active site in the PHA biosynthesis, it is instructive to compare the geometry of the catalytic triad in CoA-bound fully open form PhaC$_{Cs}$-CAT (3.10 Å) to the highest resolution semi-closed form (1.48 Å) (**Fig. 18a**). Although catalytic Cys291 and Asp447/Asn447 from both structures are well overlapped, His477 is shifted out from the catalytic pocket in the fully open form structure. This shift is a result of large unfolding of the segment between β9 and β10 strands, containing β9-α5 loop, short α5 helix and α5-β10 loop (α5-β10 loop is not defined in the current structure because of its poor electron density). In previous studies, His477 was proposed to activate Cys291 by accelerating the deprotonation of Cys thiol group, but not involved in the elongation

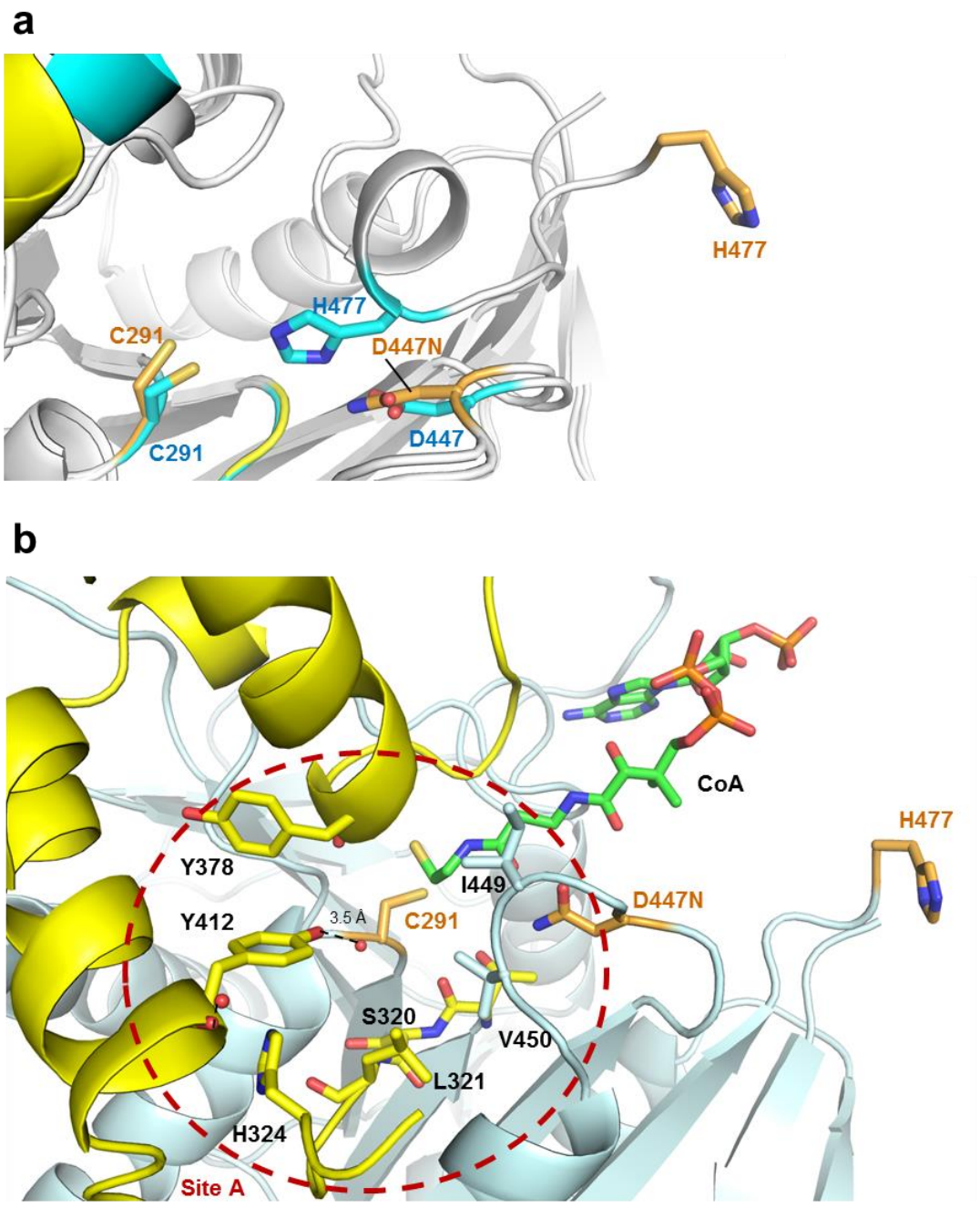


Figure 18. Comparison of the catalytic triad

(**a**) Overlay of catalytic triads from open form (yellow) and semi-closed (cyan) form of $PhaC_{Cs}$-CAT. While C291 and D447/D447N are well overlap, His477 is shifted out from the catalytic site. CoA was not shown in this figure.

(**b**) A conserved water molecule was observed in the possible binding pocket Site A interacts with Y412 at a distance of 3.5 Å. His477 shifted from 3.8 Å (semi-closed form) to ~22 Å (fully open form) away from Cys291.

process of PHA[35]. The fact of His477 movement out from the catalytic pocket without disrupting the core arrangement of the catalytic site suggests that this structure could represent an elongation state of PhaC.

Another catalytic residue, Asp447 was proposed to act on the PHA elongation process by attacking the hydroxyl group in the acyl moiety of acyl-CoA substrate[35]. In our complex structure, hydrogen bonds between the side-chain of Asn447 (Asp in the wild-type) and O_5 of CoA was observed. This interaction was supported by previous reports that D480N mutation in $(His)_6$-tagged PhaC of *C. necator* decreased the reaction rate to 0.0008 units/mg from 20 unit/mg observed in wild-type[68].

Site A in the open form structure consists side-chains of His324, Tyr412, Ser320, Leu321, Ile449 and Val450, as found in the closed form (**Fig. 10b**). In addition to these residues, however, the main-chain of Tyr378 also participates in forming Site A of the open form. A water molecule was found to interact with the hydroxyl group of Tyr412 at a distance of 3.5 Å (**Fig. 18b**). This observation suggests Site A might be the possible binding pocket for the acyl moiety from acyl-CoA. Contrary to Site A showing conservation of residues forming the site, Site B (**Fig. 10b**) is open up with the polar residue Asn220 flipped out from the cavity and facing outward in the open form structure.

3.2.3 CAP subdomain without dissociation of its dimeric form.

Structural comparison of the dimeric arrangement of our free and CoA-bound forms of $PhaC_{Cs}$-CAT and the reported free form of $PhaC_{Cn}$-CAT (PDB 5HZ2) revealed three different dimeric arrangements, represent a semi-closed homodimer, a fully closed-open heterodimer, and a partially open homodimer, respectively

(**Fig. 19a-d**). When the mol A of all three structures were superposed, mol B of CoA-bound PhaC$_{Cs}$-CAT was shifted toward one side (upward in **Fig. 19a**), while mol B of PhaC$_{Cn}$-CAT were shifted toward the opposite side (downward in **Fig. 19a**). These differences in dimeric arrangements suggest that the CAP subdomain is conformationally dynamic rather than stable and static. Since all the core subdomain structures are relatively similar in these structures, the dynamic properties of the CAP subdomain conformation might be the key factor in the PHA synthase activity.

Comparison of the N-terminal end of the catalytic domain reveals that the positions and distances are different between the closed and open forms. The N-terminal end of the semi-closed homodimer (35.2 Å, αC atoms of Ile177 between mol A and B) is closer to each other (**Fig. 19b**) than the fully closed-open dimer (46.6 Å, αC atoms of Ile177 between mol A and B) or partially open dimer (54.6 Å, αC of Ala202 between mol A and B) (**Fig. 19c, d**). These observations suggest that the N-terminal domain not only responsible for dimerization but also involved in regulating the closed-open conformational equilibrium, which should be important in the initiation and elongation steps of the PHA polymerization catalysis.

3.2.4 Helices αA and αB are the key in open-closed conformational change.

To understand the dynamic conformational transition, the structural differences between the fully closed and fully open forms of PhaC$_{Cs}$-CAT are compared. The overlay of these structures showed a highly comparable conformations of the α/β core subdomain and a largely deviated conformations of the CAP subdomain (**Fig. 20a, b**).

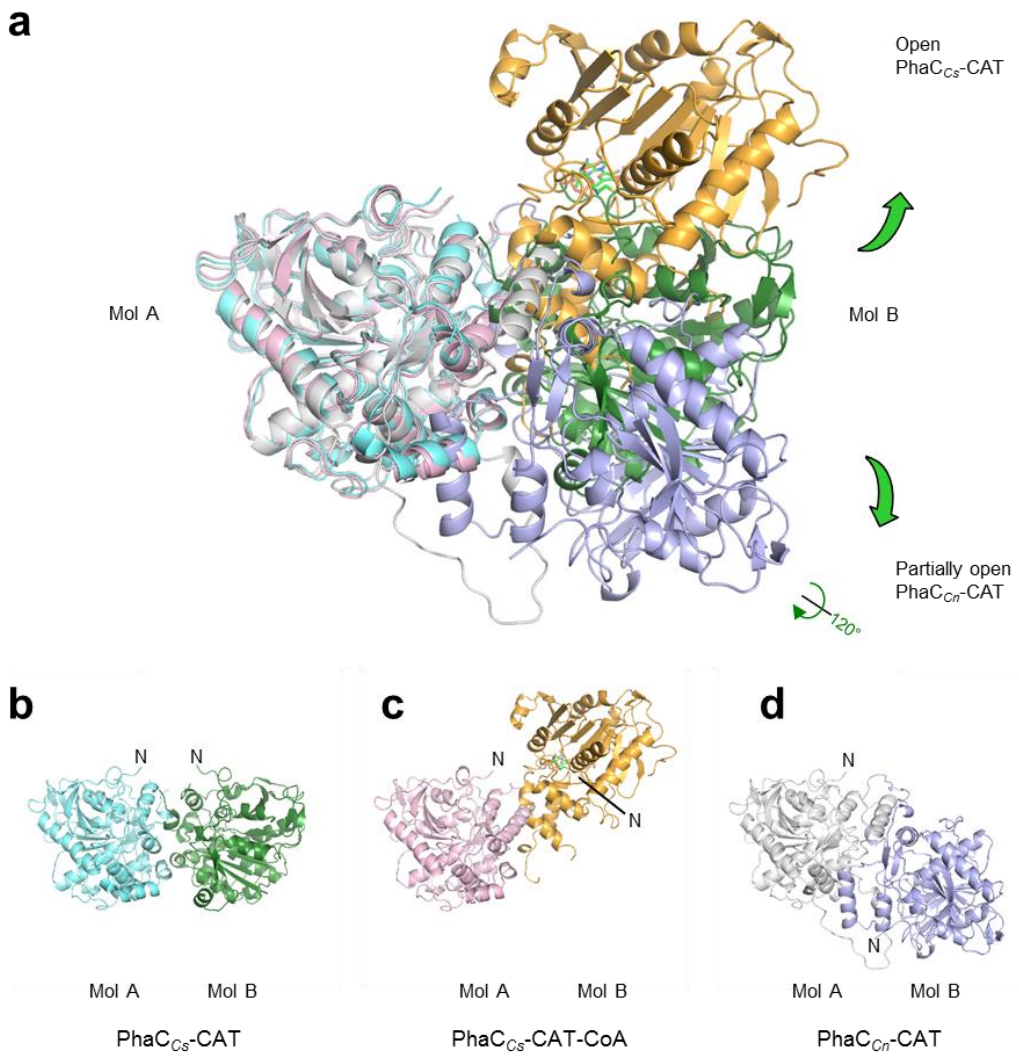


Figure 19. Conformational changes in the CAP subdomains in the dimeric forms without dissociation.

(**a**) Structural overlay of the closed-closed PhaC$_{Cs}$-CAT homodimer (PDB 5XAV), the closed-open PhaC$_{Cs}$-CAT+CoA heterodimer, and the partially open-open PhaC$_{Cn}$-CAT homodimer (PDB 5ZH2).
(**b**) The closed-closed PhaC$_{Cs}$-CAT homodimer. The N-terminal end distance is 35.2 Å.
(**c**) The closed-open PhaC$_{Cs}$-CAT+CoA heterodimer.
(**d**) The partially open-open PhaC$_{Cn}$-CAT homodimer (PDB 5HZ2).

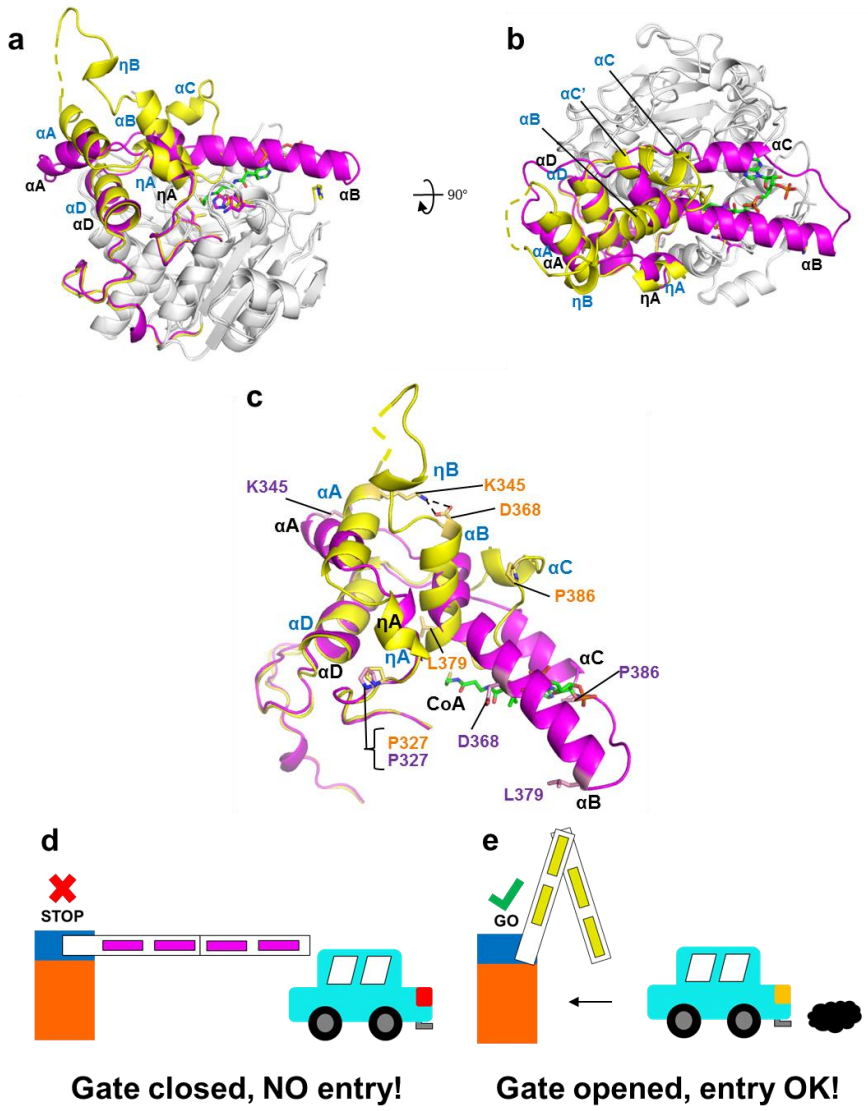


Figure 20. Dynamic conformational changes at CAP subdomain regulate closed-open form for substrate entry.

(**a**) Overlay structures of fully closed and open form PhaC_{Cs}-CAT (D447N). The α/β core subdomain displays high similarity, while the CAP subdomain undergoes significant conformational changes to allow substrate entry. The closed form contains the α/β core subdomain (pink), CAP subdomain (magenta); The open form, the α/β core subdomain (pale cyan), CAP subdomain (yellow). Coenzyme A is shown in sticks form (atom color codes are C in green, N in blue, O in red and S in yellow)

(**b**) As in a, but a top view down to the active site.

(**c**) The LID region (P327-P386) is complete restructured with the first helix movement was observed at K345, which appear to be interacting with D368 in the open form. In the closed form, a long helix αB consisting D368-L379 covers the catalytic pocket. The residue P386 also move in. The color codes are magenta for the fully closed form and yellow for the fully open form.

(**d, e**) Figurative presentation of the transition from the closed to open form using 'boom gate' analogy. (**d**) In the closed form, the gate arm (long αB helix) blocks the car (substrate) entry. (**e**) In the open form, the gate arm (long αB helix) is lift up and retracted and allows the car (substrate) to enter.

The restructured region is not limited to LID region (residues Pro327–Pro386) mentioned in **Fig. 11c**, but also includes the class I, II-conserved helix αC (Phe387-Asn394).

Comparison of the CAP subdomain between fully open (yellow) and fully closed (magenta) forms provide informative insights into the conformational transition (**Fig. 20c**). The major differences are observed at the helix αB and αC, which are responsible for blocking the substrate entry in the fully closed form. In the overlay structures, CoA intercepts with helices αB and αC of the fully closed form (**Fig.20c**). This observation suggests αB and αC helices are flexible which is also supported by the fact that part of these helices was disordered in the semi-closed form structure. The movement of helices αB and αC resembles a retracting mechanism of 'Boom gate', which covered the catalytic site and blocks the substrate entry (**Fig. 20d, e**).

Based on structural comparison of each segments of the CAP subdomain, the structural transition is mediated by three α-helices through the interactions with residues from the α/β core subdomain as well as nearby helices αA and αD (**Fig. 21a-h**). A rigid movement of αA was observed by tilting the side-chain of Arg409 (helix αD), acting like a hinge, interacting with Glu336 and Glu343 (helix αA) (**Fig. 21a, b**). This rearrangement of hydrogen bonding interactions induces a movement tilting up the C-terminal end of helix αA to create a space required for accommodation of retracted helix αB.

The extended helix αB of the fully closed form is the main region blocking the substrate from entering into the catalytic site. Several interesting intramolecular interactions are observed in this region. In the closed form, Asp368 (αB helix) interacts with His448 (next to catalytic Asp447) (**Fig. 21c**). This interaction links

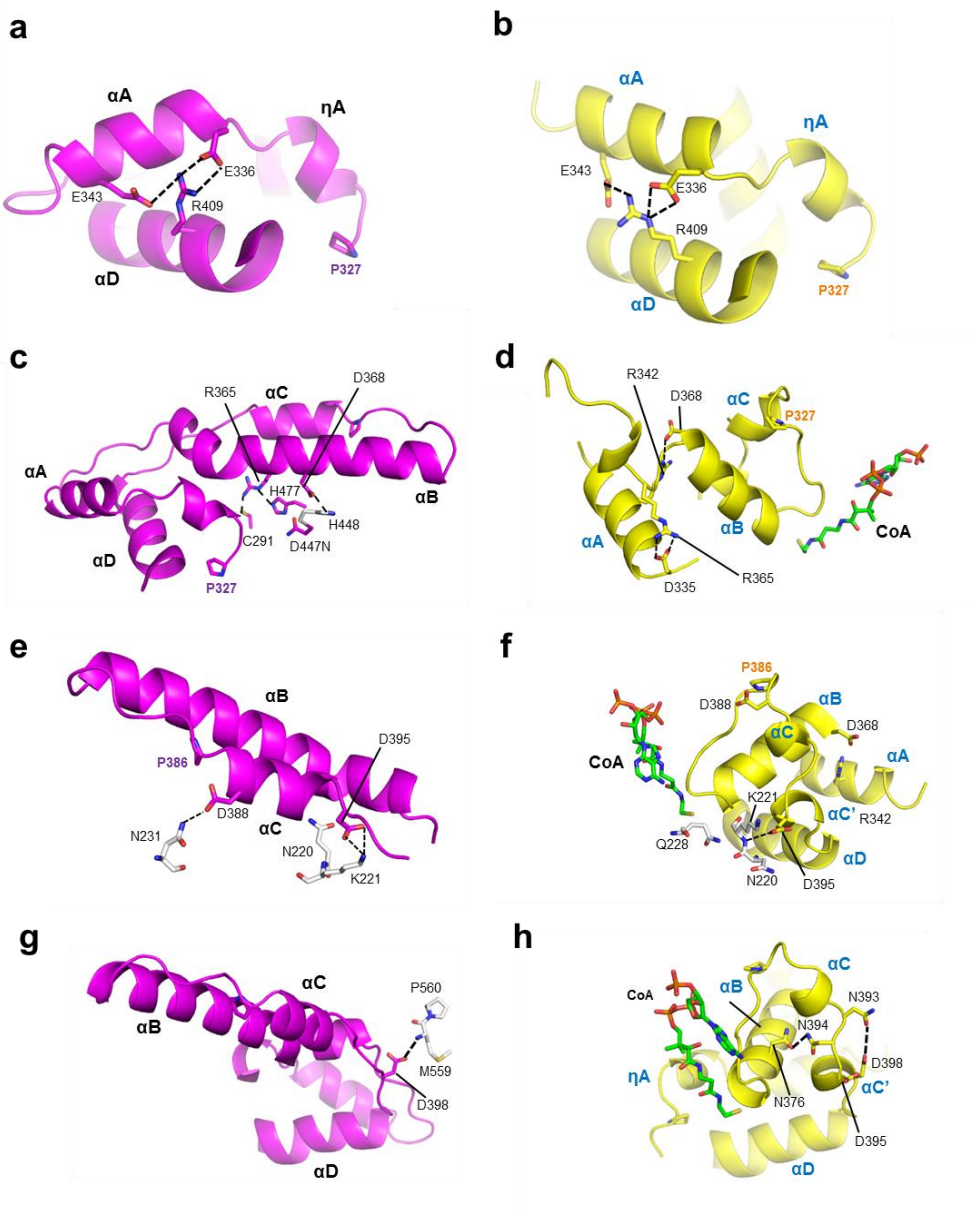


Figure 21. Helices αB and αC are the key in the open-closed conformational change.

(**a**), (**c**), (**e**) and (**g**) show LID segments of the CAP subdomain from the closed form (magenta) of $PhaC_{Cs}$-CAT and (**b**), (**d**), (**f**) and (**h**) from the open form (yellow). Residues from the α/β core domain are depicted in grey.

(**a**), (**b**) The first hinge residue R409 interacts with E336 and E343, for tilting up αA helix.

(**c**), (**d**) R365 and D368 hold αB helix closer to the catalytic triad in the closed form and retracts αB helix in the open form.

(**e**), (**f**) The second hinge D395 interacts with K221. In the open form, D388 being free from interaction with N231 and contributes to the opening.

(**g**), (**h**) D398 contributes to lift up αC helix in the open form, in order to open up a substrate entry pathway by dividing αC helix.

αB helix closer to the catalytic pocket. Possibly, this is one of the key residues in releasing αB helix from the closed conformation during catalytic reaction. Another residue, Arg365 (αB helix) is positioned in the catalytic pocket and forms hydrogen bonds with Cys291 and His477. In the open conformation, both Asp368 and Arg365 from helix αB are shifted and interact with Arg342 (αA helix) and Asp355 (αA-ηA loop), respectively (**Fig. 21d**). This dynamic movement of αB helix creates a cleft at the molecular surface for CoA entry.

Beside long αB helix, αC helix is required to cover the catalytic site completely. This helix is stabilized by Asp388 and Asp395, which interact with the side chain of Asn231 and the side chain of Lys221, respectively (**Fig. 21e**). In the open form, αC helix (Phe387-Asn394 in the closed form) was divided into two shorter helices, namely αC (Leu389-Trp392) and αC′ (Asp395-Asp398) helices. In this transition, Asp388 moves by being free from its interaction with Asn231 and contributes to the opening (**Fig. 21f**). While Asp395 still holds onto the main-chain of Lys221, this residue acts as a second hinge.

The other important structural rearrangement is observed at the interaction of Asp398 (αC-αD loop) with the main-chain N of Met559 from the C-terminal end in the closed form (**Fig. 21g**). In the open form, however, Asp398 interacts with the main-chain O of Asp395 to hold part of helix αC in a compact form and contributes to opening and the CoA entry (**Fig. 21h**).

CHAPTER 4 DISCUSSION

I have presented a high resolution structure of PhaC$_{Cs}$-CAT. To my surprise the structure of the CAP subdomain was found to differ from that of the recently reported structures of PhaC$_{Cn}$-CAT, another class I PHA synthase[69,70]. Inspection of the PhaC$_{Cs}$-CAT and PhaC$_{Cn}$-CAT structures suggests that the catalytic domain of PHA synthases exists in an open-closed equilibrium in solution by means of conformational changes in the CAP subdomain (**Fig. 22a, b**). The two reported structures of PhaC$_{Cn}$-CAT were obtained from essentially the same crystals produced under similar crystallization conditions using ammonium sulfate at pH 7. These crystals were obtained from crystallization of the full-length protein and found to have a C-terminal catalytic domain by partial degradation during crystallization. Although one of the reported structures is a mutant PhaC$_{Cn}$-CAT protein comprising a C319A mutation on the catalytic Cys residue of PhaC$_{Cn}$[69] and displayed a disordered D-loop, the two reported structures display similar partially open form structures, which may be stabilized by an artificial disulfide bond formed between Cys328 and Cys438 of PhaC$_{Cn}$-CAT.

I also have determined the complex structure of PhaC$_{Cs}$-CAT+CoA. The structure adopted a dimeric arrangement (a closed-open heterodimer) consists of two protomers displaying different forms: a fully closed and a CoA-bound fully open forms (**Fig. 15a)**. Comparison of these two forms revealed an interesting conformational transition involving αB and αC helices. The conformational transition is similar to a retracting mechanism of a 'boom gate' (**Fig. 20d, e**). A CoA molecule was found to be located at the active site and its binding mode and the interactions

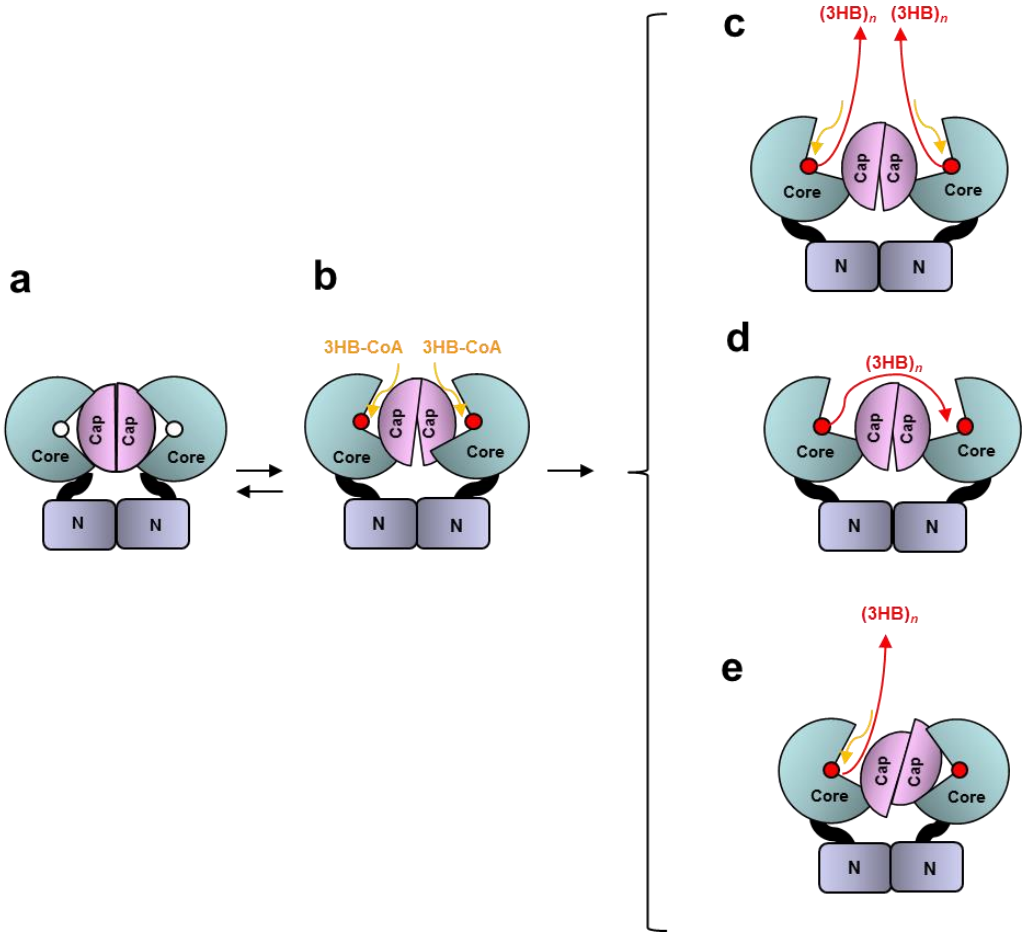


Figure 22. A model of PhaC activation in dimeric form.

(**a**) The closed form as observed in our PhaC$_{Cs}$-CAT structure. The N domain is proposed to contribute to stabilization of the dimer.

(**b**) The partially open form as observed in the PhaC$_{Cn}$-CAT structure. The structure provides a possible path to the active site, but no apparent exit path for product.

(**c**) The single active site provides a full active site architecture for initiation of acylation and chain elongation.

(**d**) The Cys-bound product in one protomer attacks the 3HB-Cys thioester of the other protomer in the dimer for chain elongation.

(**e**) Only single active site is required for full activity. One protomer in the open form catalyzes PHA polymerization and the other protomer in the closed form stabilizes the open form by directly interacting the CAP subdomain.

mediating the binding is discussed. Interestingly, the disordered region found in the free form of PhaC$_{Cs}$-CAT was found to interact with CoA, suggesting that this flexible region could play a role of CoA sensing and a trigger of the dynamic conformational changes in the CAP subdomain for full binding of CoA to the active site.

The active site architecture of the conserved catalytic triad (Cys-His-Asp) of the semi-closed form of PhaC$_{Cs}$-CAT suggests that His477 is located at a position where the side-chain imidazole ring could accelerate deprotonation of the side-chain thiol group of the active center residue Cys291 to trigger nucleophilic attack of the thiol group to the thioester of acyl-CoA, which would yield the acyl-Cys291 intermediate. In the semi-closed form structure, negatively charged Asp447 assists to enhance the basicity of His477 by formation of a direct hydrogen bond and also enables His477 to access to Cys291 by fixing the side-chain conformation of His477 (**Fig. 9a**). The side-chain conformation of Asp447 is tightly restricted and stabilized by forming multiple hydrogen bonds to the main-chain amide groups of β8-α4 loop (Asp447-His448-Ile449) (**Fig. 9a**). These interactions are retained in the open form, and catalytic center Cys291 and Asn447 retain their conformations and relative positions inside the catalytic site, whereas His477 is shifted out from the catalytic site by large conformational changes in the segment between β9 and β10 strands with unfolding of α5-β10 loop. The fact that the His477 flip out without destruction of the structural integrity of the catalytic site suggests that His477 and the His477-containing segment between β9 and β10 should be very flexible and may imply that His477 has a role in initiation by accelerating deprotonation of Cys291, but not in elongation in the polymerization of acyl moieties. Previous reports suggest that

the Asp residue of the catalytic triad acts as a general base catalyst to accelerate deprotonation of the 3-hydroxyl group of HB in the step involving elongation of the PHA product[38,68]. In the open form, Asn447 (Asp in the wild-type) stabilizes the CoA-binding by directly interacting with the pantetheine arm of the bound CoA molecule (**Fig. 17**). Unfortunately, the precise geometry of the substrate-binding mode and the role of the Asp residue with 3-hydroxyl group of HB could not be determined from the present structures of the free and CoA-bound forms and further studies should be carried out for the clarification.

Two catalytic mechanisms for the PHA synthesis by PhaC have been proposed in the context of dimerization of the synthases[76]. One mechanism is referred to as the non-processive ping-pong model that requires two sets of active sites for PHA chain elongation with chain transfer between the two active center Cys residues across the dimer interface[77,71]. The second mechanism involves a processive model that requires a single active site for PHA chain elongation and a non-covalent intermediate, in addition to a covalent intermediate bound to the Cys residue at the active center during the catalytic cycle[38,41,49,67,78]. Our dimeric structures show that the two active sites should be too distant (28.1 Å in semi-closed form; 41.5 Å in open form) for successive chemical reactions and seems to favor the mechanism comprising the use of a single active site. A similar conclusion was also discussed previously in relation to a dimeric PhaC$_{Cn}$-CAT structure, which shows a long distance (~33 Å) between the active Cys residues[69]. In this model, the substrate enters the same substrate-binding tunnel, although chain product is elongated through another path to the protein surface. The first 3HB-CoA forms a 3HB-Cys covalent bond, and the resultant free CoA is released from the tunnel (**Fig. 23a**). The second 3HB-CoA enters the tunnel

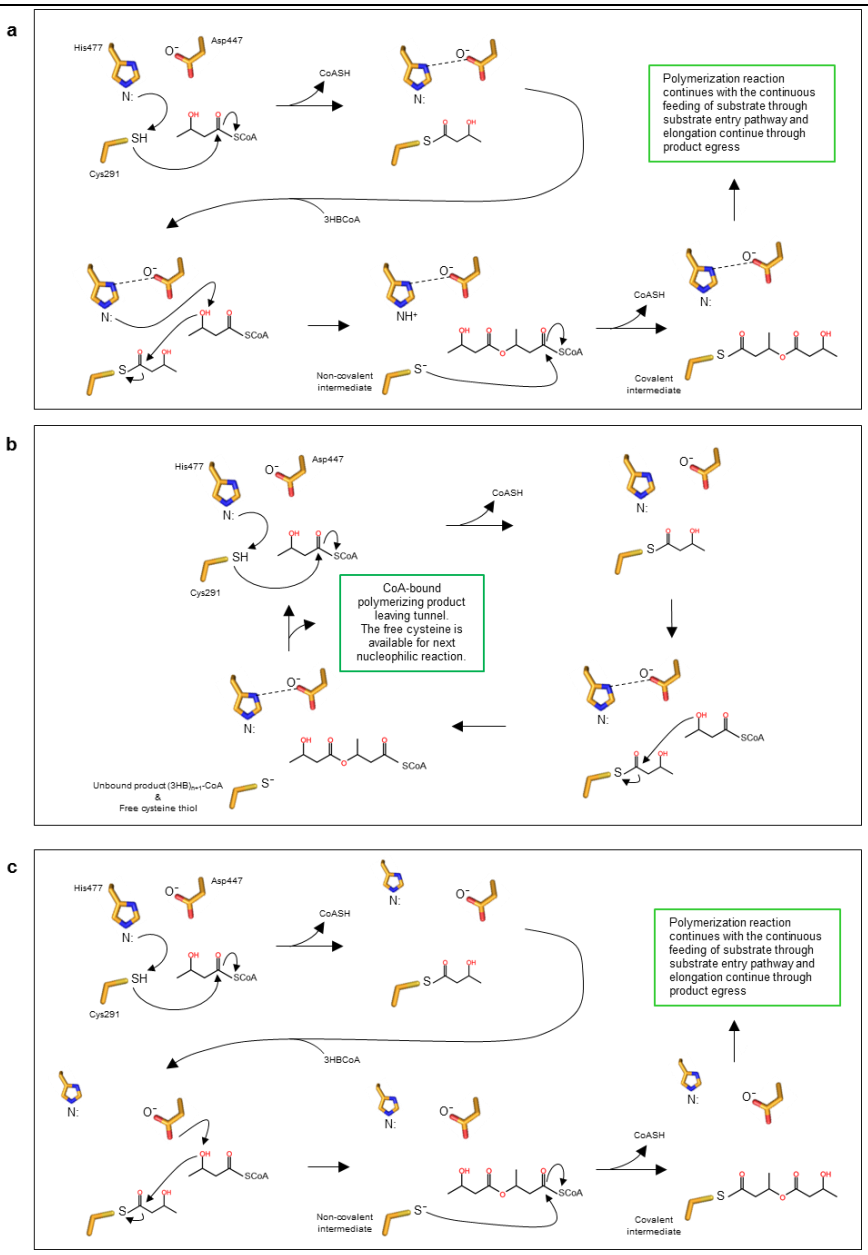


Figure 23. Proposed mechanism of PHA synthase.

(**a**) Processive single active site model (In and out tunnels) which is also proposed in reference[69]. This model requires a single active site for PHA chain elongation and a non-covalent intermediate, in addition to a covalent intermediate bound to the Cys residue at the active center during the catalytic cycle.

(**b**) Alternative processive single active site model (Single tunnel) which is proposed in reference[70]. In this model, two substrates share the same substrate-binding tunnel and the first 3HB-CoA produces 3HB-Cys. The second 3HB-CoA attacks 3HB-Cys to produce $(3HB)_2$-CoA, which is released from the active site. The cycle is repeated with newly entered 3HB-CoA to produce 3HB-Cys, and then the following $(3HB)_2$-CoA enters the active site to produce $(3HB)_3$-CoA, which is again released from the active site.

(**c**) Similar to mechanism **a**. The His477 involved in the initiation but not elongation.

and attacks the 3HB-Cys thioester bond with the hydroxyl group of the 3HB unit to produce a $(3HB)_2$-CoA intermediate and free the Cys residue of the active center. This Cys residue is available once again to attack the thioester bond of the $(3HB)_2$-CoA intermediate to produce $(3HB)_2$ covalently bound to the Cys residue and release of free CoA from the active site. This cycle is repeated with newly entered 3HB-CoA to produce $(3HB)_{n+1}$ covalently bound to the Cys residue (**Fig. 22c**). Thus, the growing 3HB polymer is bound to the enzyme at the end of each cycle, which is consistent with the fact that covalent catalysis occurs using the saturated trimer CoA (sT-CoA), an analogue of $(3HB)_3$-CoA in which the terminal hydroxyl group is substituted with a hydrogen blocking the attack by the terminal hydroxyl group[38,48,67,71.]

A slightly different model has also been proposed with essentially the same $PhaC_{Cn}$-CAT dimeric structure[70] (**Fig. 23b**). In this alternative model, two substrates share the same substrate-binding tunnel and the first 3HB-CoA produces 3HB-Cys as in the aforementioned model. The second 3HB-CoA attacks 3HB-Cys to produce $(3HB)_2$-CoA, which is released from the active site. This cycle is repeated with newly entered 3HB-CoA to produce 3HB-Cys, and then the following $(3HB)_2$-CoA enters the active site to produce $(3HB)_3$-CoA, which is again released from the active site. This model is inconsistent with the presence of a stable product bound to the active Cys residue[71]. If the $(3HB)_3$-CoA that is produced is held in the active site and attacked by the active Cys residue again to produce $(3HB)_3$-Cys, chain elongation would then require an inter-subunit reaction (**Fig. 22d**).

The proposed ping-pong mechanism requires two thiol groups located at a distance short enough to shuttle back and forth the growing $(3HB)_n$ chain between the two thiols. This may be possible if there is another Cys residue near active site

Cys291. We found a Cys residue (Cys218) that was located in β4-η1 loop near Site A at the active site (**Fig. 24a, b**). Although Cys218 is conserved in PhaC$_{Cn}$ but is not a universally conserved residue, Cys218 may assist in the catalytic action of Cys291 provided that some conformational changes are induced around the active site. However, our mutations involving replacement of Cys218 with Ala or Ser residues resulted in no significant effect on the activity (**Fig. 24c**). Another interesting residue is Thr319 of β7-ηA loop that forms a hydrogen bond (3.5 Å) to Cys291 and is conserved in Class I and IV synthases. Replacement of Thr319 with Ala or Ser, however, had no significant effect on the activity.

In the PhaC$_{Cs}$-CAT+CoA complex, the dimer consists of one CoA-bound fully open form stabilized by a fully closed form through direct interactions involving the CAP subdomain (**Fig. 15a**). Together with the fact that His477 is shifted out from the active site, a dimer activation mode is proposed (**Fig. 22e**). In this mode, only one protomer is active and polymerizes PHA. The reaction mechanism is similar to the processive model with a single active site but His477 only involved in initiation (**Fig. 23c**). The catalysis of the elongation of 3HB moieties involves Asp447-mediated activation of the hydroxyl group from the second 3HB-CoA molecule and this cycle is repeated with newly entered 3HB-CoA to produce (3HB)$_{n+1}$ covalently bound to the Cys residue[35].

A previous mutational study of PhaC$_{Cs}$ reported that the Class I-conserved residue Ala479 represents at least one of the critical residue required for substrate specificity, as determined by various site-specific mutational assays both *in vivo* and *in vitro*, and production tests of copolymers such as P(3HB-*co*-3HHx) [79]. The greatest

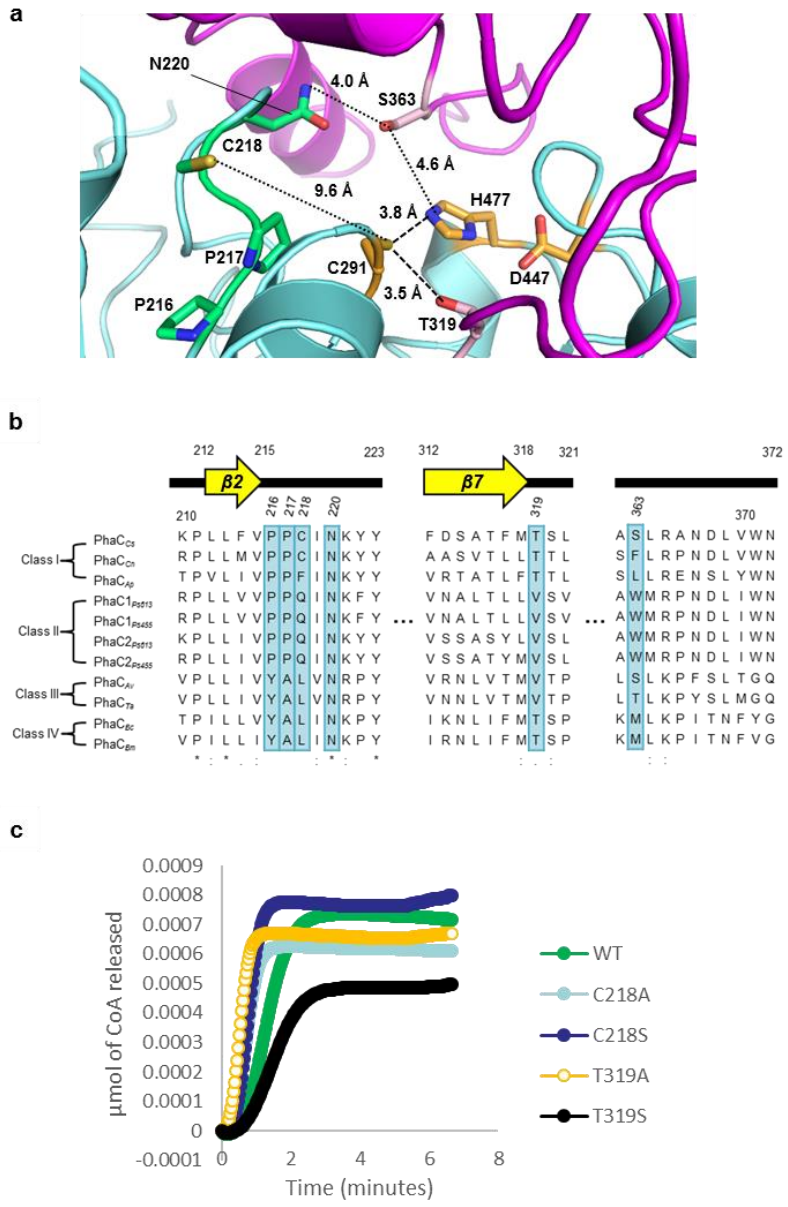


Figure 24. The active site of PhaC$_{Cs}$-CAT contains residues at positions where assistance of active center Cys291 or intermediate binding may be possible at the active site.

(**a**) The active site of PhaC$_{Cs}$-CAT contains two polar residues, and Cys218, which may act to stabilize intermediate binding, and Thr319 may assist with activation of active center Cys291. The active site of the closed form of PhaC$_{Cs}$-CAT also contains Ser363 from the LID region of the CAP subdomain. Hydrogen bonds (broken lines) and interesting distances (dotted lines) are indicated.
(**b**) Partial alignment of the sequence showing Cys218, Thr319 and Ser363 of PhaC$_{Cs}$-CAT.
(**c**) Enzymatic activity of wild-type and mutant PhaC$_{Cs}$ with substrate 3HB-CoA. Released CoA was monitored.

enhancement in PHA biosynthesis was observed with the A479G mutation which resulted in a 1.6-fold increase, and mutations of large nonpolar residues (Met, Trp and Val) also enhanced the activity. In the semi-closed form structure, Ala479 is located within α5 helix and the side chain protrudes into a depression of the molecular surface formed by loops (β4-α1, β9-α5 and α5-β10 loops) from the core subdomain, and is partially covered the 3_{10}-helix ηB' and the following loop of the LID region from the CAP subdomain (**Fig. 25a**). The side chain of Asp368 of ηB'-helix forms two hydrogen bonds to the main chains of α5 helix. These interactions are typical of N-terminal caps stabilizing the helix, although in this case crowding is induced by the presence of ηB'-helix and the following Val370 residue. Therefore, it was expected that replacement of Ala479 with residues having large side chains would result in the loss of these direct interactions between α5 helix and ηB'-helix. We speculate that this loss results in weakening of the interactions between the LID region and the core subdomain, and stabilizes the active form of this enzyme by releasing the LID region from the active site. Interestingly, a maximum 4-fold increase in 3HHx incorporation was observed for the A479S mutation, while the A479T mutation showed ~2.8-fold enhancement[79]. Since Ala479 is surrounded by polar residues (Ser475 and Arg490), it was reasonable to postulate that replacement of Ala479 with Ser or Thr would facilitate hydrogen-bonding interactions with the polar residues and stabilization of α5 helix harboring the active residue His477, which may be important for enzyme activity. In contrast to these possible interactions involving Ala479 for stabilization of α5 helix, in the open form Ala479 was located at the disordered region (**Fig. 25c**), suggesting that Ala479 may assist His477 only for the initiation reaction.

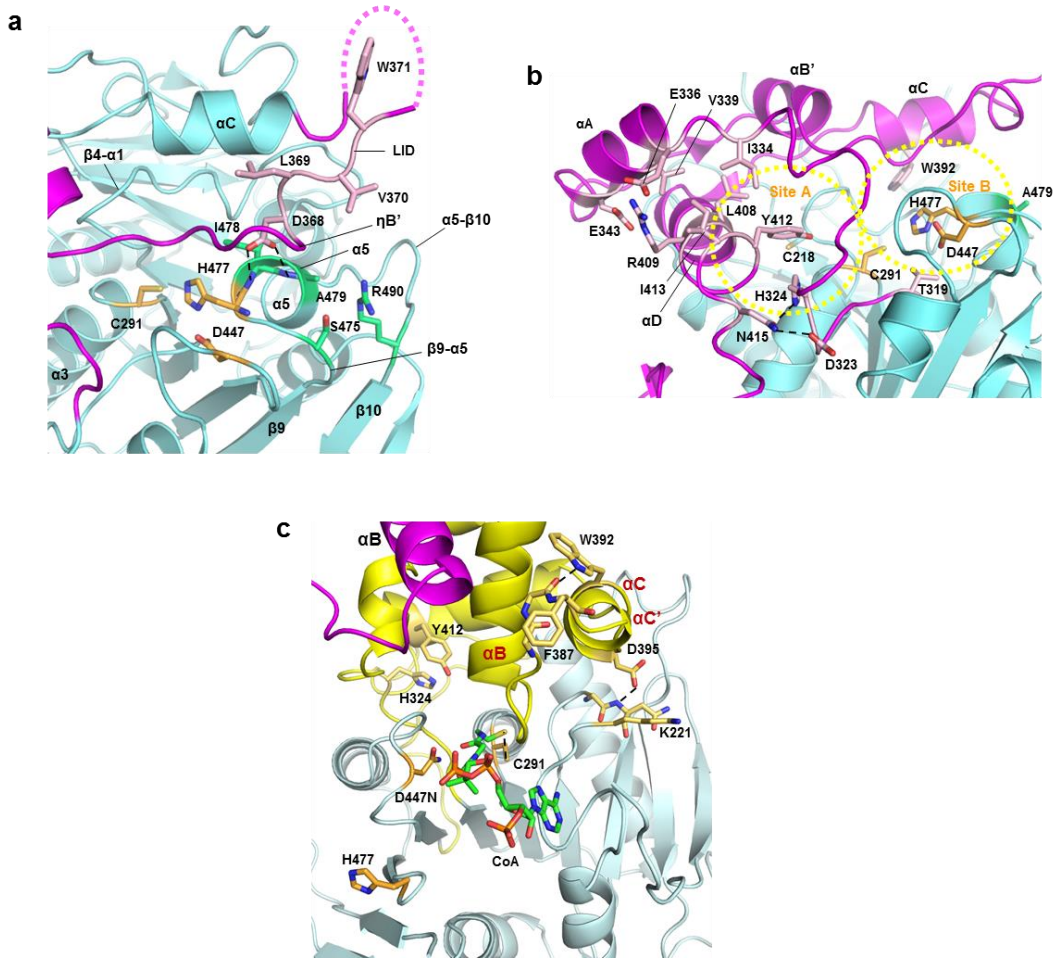


Figure 25. The channel of the active site of PhaC$_{Cs}$.

(**a**) A close-up view of the β9-α5-β10 segment forming part of the active site of the semi-closed PhaC$_{Cs}$-CAT. Ala479 of PhaC$_{Cs}$ is located at α5 helix, where His477 of the catalytic triad is also located. Asp368 of the LID region stabilizes α5 helix by forming hydrogen bonds to the main chain amide groups of Ile478 and Ala479 residues which form the helix. Ala479 corresponds to Ala510 of PhaC$_{Cn}$.

(**b**) A view of the channel at the active site of the semi-closed form of PhaC$_{Cs}$-CAT. Trp392 of PhaC$_{Cs}$ is located in the αC helix of the CAP subdomain and faces Site B of the channel where a water cluster is present. Tyr412 and Ile413 are located in the αD helix of the CAP subdomain. Tyr412 projects into Site A of the channel, while Ile413 forms a hydrophobic core with other aliphatic residues.

(**c**) A top view of the active site of the open form of PhaC$_{Cs}$-CAT. Ala479 is located at the disordered region in the structure. His477 is shifted out from the active site. The residues Trp392 and Asp395, contribute to holding αC helix to create an open space for CoA binding by forming hydrogen bonds with backbone amide groups of Pro386 and Lys221, respectively.

The CAP subdomain provides αC and αD helices as building blocks of the active site cavity filled with a cluster of water molecules. In my semi-closed form structure, the C-terminal portion of the LID region of the CAP subdomain is disordered and is followed by αC helix docked to the core subdomain. Two highly conserved residues, Trp392 and Asp395, are present in αC helix (**Fig. 7**). Asp395 stabilizes αC helix docking by forming a salt bridge with highly conserved Lys221 (β4- η1 loop) of the core subdomain. Trp392 is part of the group of residues forming the active site and stabilizes the αC helix by anchoring the nonpolar side chain to the nonpolar site (**Fig. 25b**). Since αC helix is important for dimerization, mutation of Trp392 resulted in disruption of dimerization of $PhaC_{Cn}$ (Trp425 in $PhaC_{Cn}$)[80] and abolished the activity of Class II PhaC1 from *Pseudomonas aeruginosa* (Trp398 in $PhaC1_{Pa}$)[81]. However, in the open form, Trp392 and Asp395 were involved in opening the LID region by lifting up αC helix (**Fig. 20e, f, h**). Asp395 interacts with the main-chain amide of Lys221 (3.2 Å) and plays a role of a hinge (**Fig. 25c**). Trp392 stabilizes the interaction between αC and αB helices by forming a hydrogen bond with the main chain of Pro386 and anchoring its nonpolar side chain to the hydrophobic region in dimerization (**Fig. 25c**).

At the channel of the active site, αD helix is the building block of Site A (**Fig. 10b**). Tyr412 and Ile413 of αD helix form a hydrophobic core together with other nonpolar residues from αD helix (Met407 and Leu408) and αA helix (Ile334 and Val339). Mutations of residues corresponding to Tyr412 and Ile413 in $PhaC_{Cn}$ resulted in a reduction in enzyme activity[82]. The C-terminal end of this helix possesses highly conserved Asn415, which forms hydrogen bonds with β7-ηA loop (Asp323 and His324), and contributes to formation of the Site A channel. As

described above, Tyr412 and His324 are polar residues facing the channel and may participate in substrate binding. It is noteworthy that Tyr412 is conserved in Class I, III and IV PHA synthases, which utilize SCL substrates. In contrast, Class II synthases, which utilize MCL substrates, have a preference for Phe at this position.

In the open form, the CoA binding model gives us a hint for the specific enzyme-substrate interacting mechanism. In the open form structure, Site A consists side-chains of His324, Tyr412, Ser320, Leu321, Ile449 and Val450 and the main-chain of Tyr378 (**Fig. 26a**). A conserved water molecule at Site A forms a hydrogen bond with the hydroxyl group of Tyr412. In our model, the common C4 SCL substrate, 3HB-CoA can be positioned to fit loosely in Site A (**Fig. 26b**). Interestingly, PhaC$_{Cs}$ favors a C6 substrate 3HV-CoA is able to be fitted snugly onto the active site (**Fig. 26c**). Class I PhaC$_{Cs}$ does not favor towards MCL C8 substrate 3HO-CoA and this can be explained by the steric hindrance (**Fig. 26d**). Despite the exact geometry of the binding mode is still remained to be determined, the modeling contributes to our understanding of its acyl-binding mode.

A study employing random mutagenesis revealed a beneficial mutation (F518I) of PHA synthase from *Aeromonas punctate* (PhaC$_{Ap}$)[80]. Compared with wild-type PhaC$_{Ap}$, the F518I mutation in PhaC$_{Ap}$ (equivalent to Tyr492 in PhaC$_{Cs}$) increased the relative activity to 480% in an *in vitro* synthase activity assay and 120% in terms of *in vivo* PHA accumulation ability. In our structure, Tyr492 is located within β10 strand and stabilizes β-sheet formation between β10 and β11 strands by engaging in an aromatic stacking interaction with adjacent His512 from β11 strand (**Fig. 27**). Formation of a hydrogen bond between Thr494 and His512 also contributes to stabilization of the β-sheet. Multiple sequence alignments revealed a similarity in

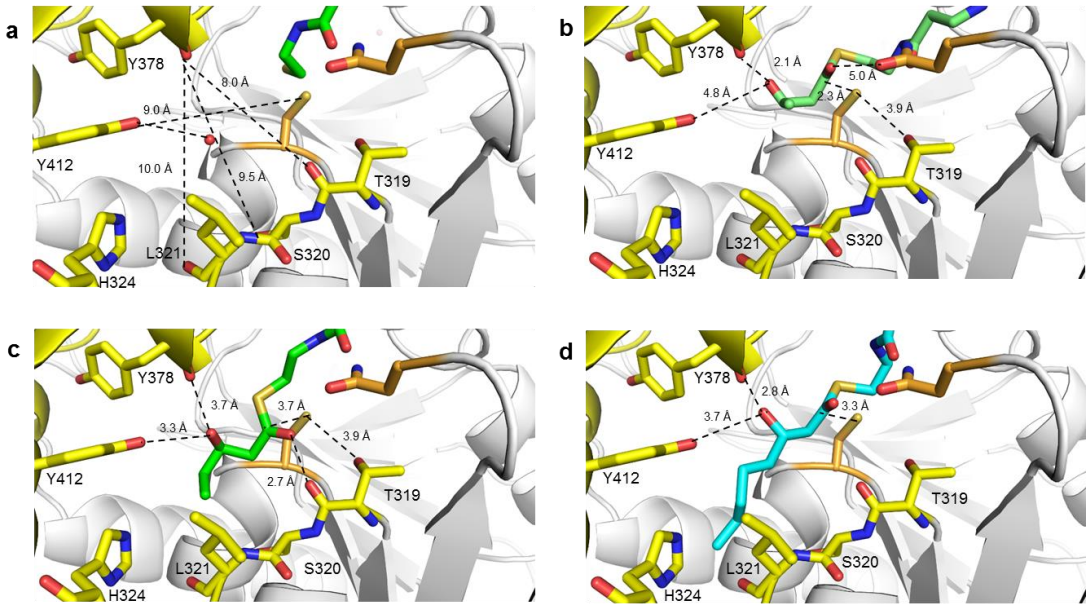


Figure 26. Modeling of acyl moieties fitting into possible binding pocket Site A.

(**a**) Original empty cavity observed at Site A in the open form. A water molecule was located at a distance of 3.5 Å to the hydroxyl group of Tyr412. Site A has estimated 9 $Å^2$ space. The CoA molecule is showed in green.

(**b**) A docking model using 3HBCoA, a common substrate. The distance of thioester carbon is at 2.3 Å from Cys291.

(**c**) A docking model with 3HVCoA, which is favored by $PhaC_{Cs}$. The 3HVCoA molecule can be fitted into the catalytic pocket.

(**d**) A docking model with 3HOCoA. In our modeling, 3HOCoA was found to be too long to fit into this pocket. This may be consistent with the fact that Class I PhaC does not favor MCL-substrates.

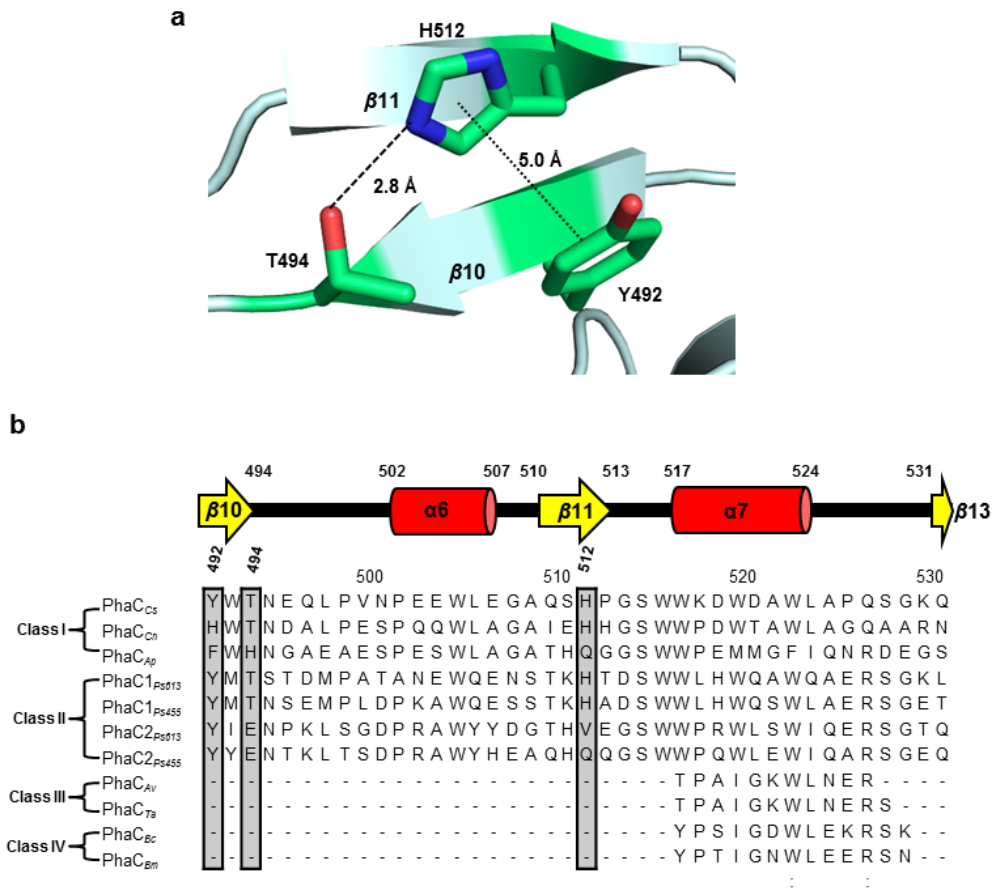


Figure 27. Conformation of the β10-β11 segment of PhaC$_{Cs}$.

(**a**) A close-up view of the β9-β10 antiparallel β-sheet, which is part of the PhaC-specific additional β9-α5-β10 segment of the core subdomain. The segment is located at the edge of the α/β core and projects from the core subdomain without making contacts with the other protomer of the closed form dimer of PhaC$_{Cs}$-CAT. Hydrogen bonds (broken lines) and the ring-ring distance (dotted line) between His512 and Tyr492 residues are indicated.

(**b**) Partial alignment of the sequence showing that Tyr492 is replaced with other aromatic residues such as Phe or His in Class I synthases, but is conserved in Class II synthases.

aromatic amino acids at the position corresponding to Tyr492 of PhaC$_{Cs}$ in Class I and II synthases. However, the residues in these positions differ between PhaC$_{Ap}$ and PhaC$_{Cs}$, where residues Tyr492, Thr494 and His512 (PhaC$_{Cs}$) are replaced with residues Phe518, His520 and Gln538 (PhaC$_{Ap}$), respectively. In PhaC$_{Ap}$, Phe518 and His520 are stacked, and Gln538 and His520 may form a hydrogen bond which stabilizes the β-sheet. The F518I mutation may improve hydrocarbon chain interactions between Gln538 and Ile518. Since catalytic residue His477 is located at the N-end of α5 helix, which is connected to strand β10, fine-tuning of this catalytic residue by the F518I mutation in PhaC$_{Ap}$ may have resulted in the enhanced activity. Further analysis should be employed to clarify the details of the mechanism of the enhanced activity associated with the F518I mutation. Comparison of the primary sequences of Class I PhaC$_{Cs}$ and Class II PhaC1 revealed the presence of conserved residues comprising tyrosine, threonine and histidine, suggesting that these PHA synthases may share the same architecture in maintaining the β strands observed in the current structure.

A previous study showed that the F420S mutation in PhaC$_{Cn}$ increased the specific activity with a significantly reduced lag phase[81]. This residue corresponds to Phe387 of PhaC$_{Cs}$, which is conserved among Class I and II PHA synthases, and is located in αC helix of the CAP domain (**Fig. 13c**). As previously discussed, Phe387 is involved in dimerization by participating in an intermolecular nonpolar interaction directly connected to the LID region, suggesting that the mutation may contribute to a reduction in the lag phase by affecting the conformational stability and/or conformational transition of the LID region. The conformational characteristics of the LID region should be important for the enzymatic activity. A previous study utilizing

random mutagenesis found that A391T and T393A mutations in PhaC$_{Cn}$ (corresponding to G358 and T360 located in αB' helix of the LID region in PhaC$_{Cs}$) resulted in synthase activity at lower efficiency[82]. In the recently reported structures of PhaC$_{Cn}$-CAT[69,70], Ala391 and Thr393 are located at the interface between the LID region and the core subdomain, and are engaged in nonpolar and polar interactions, respectively. In our PhaC$_{Cs}$-CAT structure, G358 and T360 interact with another part of the LID region rather than the core subdomain. These residues are not necessarily conserved in other synthase classes (**Fig. 7**) and may reflect the characteristic properties of each enzyme. However, the conformational properties of the LID region are affected by the component residues interacting with other parts of the enzyme and have some effect on the catalytic activity.

In conclusion, I have defined a structure of the semi-closed form of PhaC$_{Cs}$-CAT and a complex structure of the CoA-bound closed and open forms of PhaC$_{Cs}$-CAT from *Chromobacterium* sp. USM2.

In the free form PhaC$_{Cs}$-CAT structure, the catalytic site is covered by the CAP subdomain and the catalytic residues are facing a water-filled large channel inside the protein. This form displays a sharp contrast to the recently reported structure of the partially open form of the catalytic domain of PhaC$_{Cn}$. The major difference between the closed and partially open forms is found in the conformation of the CAP subdomain. Both catalytic domains of PhaC$_{Cs}$ and PhaC$_{Cn}$ form a dimer mediated by the CAP subdomain. Therefore, the arrangement of the protomers to form a dimer also differs in these two forms. These facts suggest that the CAP subdomain should undergo a conformational change during catalytic activity with rearrangement of the

dimeric form.

The complex structure of PhaC$_{Cs}$-CAT+CoA comprised of a closed-open heterodimer, which has the dimer interface formed by the CAP subdomains. The catalytic His477 residue was shifted out from the active site supporting the notion of its role in initiation. The CoA-bound form reveals the importance of CoA moiety for PHA polymerization as reported previously. The possible binding of the acyl moieties of the substrates were modeled and the structural fitting was explained. Although both closed and open forms of PhaC$_{Cs}$-CAT suggest the intrinsic conformational flexibility of the CAP subdomain without dissociation of the dimer, the role of the N-terminal domain, which is absent in our structures, in catalytic reaction was still unexplained. Also, the exact catalytic mechanism is still not fully understood without the structures of PhaC-substrate intermediate complex and full-length PhaC. Nevertheless, detailed information concerning the three-dimensional structure of the catalytic domain provides valuable clues to delineate the obscure catalytic action involved, and should assist in efforts to improve the functionality and activity of this industrially important enzyme.

ACKNOWLEDGMENT

Though only my name appears on the cover of this thesis, it would not be possible without the guidance and support from a great number of people.

Firstly, I would like to extend my sincere gratitude to supervisor, Dr. Toshio Hakoshima. My PhD journey has been an exciting and fruitful one, though, with some bitter moments. His wisdom, patience, perseverance, and guidance have inspired me to be a better scientist and could outgrow my limits.

I want to thank my advisors, Dr. Shiro Suetsugu and Dr. Tomoya Tsukazaki for their useful advices and suggestions throughout my studies.

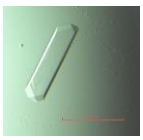
Four years in NAIST, I have learnt a lot from Dr. Kim Sun-Yong and grateful for his advices, which are the main reasons I can pick up X-ray crystallography faster. I want to credit Dr. Tomoyuki Mori for kind help in SAD, useful advices, and supports, thank you for making the lab more happening. I want to thank Dr. Yoshinori Hirano, although we have less chance to work together, his advices are always insightful. I also want to thank Ms. Shirakawa for horning up my cloning techniques and for cloning.

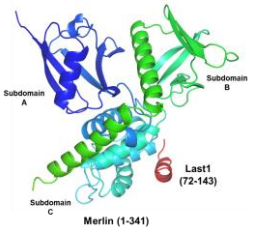
I want to acknowledge MEXT scholarship for my financial supports and NAIST for the daily life supports. To the international committee of NAIST-BIO, especially Dr. Satoko Maki, for all the efforts in establishing the international community in NAIST-BIO.

I want to memory my first protein crystal structure of Merlin-Lats1 (2.1 Å), this makes me a better 'beginner' in structural biology. Unfortunately, I cannot make it my public baby as someone else did.

My sincere thanks to my parents, Terry Chek and Mary Lim, who endlessly supported me. And to my sister, Janet, for taking care of our parents.

Lastly, thank you to those I could not name here.





The synchrotron radiation experiments were performed at BL41XU and BL44XU in SPring-8 with the approval of the Japan Synchrotron Radiation Research Institute (JASRI) (proposal nos. 2012A1058, 2012A6738, 2012B1205, 2012B6738, 2013A6844, 2013B1288, 2013B6844, 2014A1283, 2014A6944, 2014B6944, 2014B1533, 2015A1086, 2015B2086, 2015A6549, 2015B6549, 2016A6648, 2016B6648, 2016A2510, 2016B2510, 2016A2519, 2016B2519, 2017A2502, 2017A6759 and 2017B6759) and were also performed at BL-1A, AR-NE3A and AR-NW12A in the Photon Factory (PF), High Energy Accelerator Research Organization (KEK) in Japan (Proposal No. 2014G706).

REFERENCES

1. North, E.J., and Halden, R.U. Plastics and environmental health: The road ahead. *Rev. Environ. Health.* **28**(1), 1-8 (2013).

2. Halden, R.U. Plastics and health risks. *Annu. Rev. Public Health*. **31**, 179-194 (2010).

3. Plastics - the facts 2016" (PlasticsEurope, Brussels, Belgium. 2016); http://www.plasticseurope.org/Document/plastics---the-facts-2016-15787.aspx?Page=DOCUMENT&FolID=2.

4. Hopewell, J., Dvorak, R., Kosior, E. Plastics recycling: challenges and opportunities. *Phil. Trans. R. Soc. B.* **364**, 2115-2126 (2009).

5. Federal Register: The Daily Journal of the United States Government. Indirect Food Additives: Polymers. Available at: https://www.federalregister.gov/documents/2012/07/17/2012-17366/indirect-food-additives-polymers

6. European Food Safety Authority. Bisphenol A. Available at: http://www.efsa.europa.eu/en/topics/topic/bisphenol

7. Singh, S., Li, SS. Bisphenol A and phthalates exhibit similar toxicogenomics and health effects. *Gene* **494**(1), 85-91 (2012).

8. Jambeck, J.R., Geyer, R., Wilcox, C., Siegler, T.R., Perryman, M., Andrady, A., Narayan, R., Law, K.L. Plastic waste inputs from land into the ocean. *Science* **347**(6223), 768-771 (2015).

9. Cózar, A. *et. al.* Plastic debris in the open ocean. *Proc. Natl. Acad. Sci.* **111**(28), 10239-10244 (2014).

10. Thompson, R.C., Moore, C.J., vom Saal, F.S., Swan, S.H. Plastics, the environmenta and human health: current consensus and future trends. *Phil. Trans. R. Soc. B.* **364**, 2153-2166 (2009).

11. Goldstein, M.C. and Goodwin, D.S. Gooseneck barnacles (Lepas spp.) ingest microplastic debris in the North Pacific Subtropical Gyre. *PeerJ*. DOI10.7717/peerj.184 (2013).

12. Laist, D.W. in Marine Debris: Sources, Impacts, and Solutions (eds. Coe, J. M. &

Rogers, D. B.) 99–139 (Springer New York, 1997). doi:10.1007/978-1-4613-8486-1_10

13. Sudesh, K., Iwata T. Sustainability of biobased and biodegradable plastics. *Clean*. **36** 433-442 (2008).

14. Hoover, R., Huughes, T., Chung, H.J., Liu, Q. Composition, molecular structure, properties, and modification of pulse starches: A review. *Food Res. Int.* **43** 399-413 (2010).

15. Jem K.J., van der Pol J.F., de Vos S. Microbial lactic acid, its polymer poly(lactic acid), and their industrial applications. In: Chen G-Q (ed) Plastics from bacteria: natural function and applications, vol. 14., *Microbiology Monographs Springer*, Berlin, pp. 323-346.

16. Steinbüchel, A. Non-biodegradable biopolymers from renewable resources: perspectives and impacts. *Curr. Opin. Biotechnol.* **16**, 607-613 (2005).

17. Chen, G-Q. and Patel, M. K. Plastics derived from biological sources: Present and future: A technical and environmental review. *Chem. Rev.* **112** 2082-2099 (2012).

18. Madison, L. and Huisman G. Metabolic engineering of poly(3-hydroxyalkanoates): From DNA to plastic. *Microbial. Mol. Biol. Rev.* **63**, 21-53 (1999).

19. Qin, L-F., Gao, X., Liu, Q., Wu, Q., Chen, G-Q. Biosynthesis of polyhydroxyalkanoate copolyesters by *Aeromonas hydrophila* mutant expressing a low-substrate-specificity PHA synthase $PhaC2_{Ps}$. *Biochem. Eng. J.* **37**, 144-150 (2007).

20. Reddy, CSK., Ghai, R., Rashmi, K.V.C. Polyhydroxyalkanotes: an overview. *Bioresour. Technol.* **87**, 137-146 (2003).

21. Steinbüchel, A., Lütke-Eversloh, T. Metabolic engineering and pathway construction for biotechnological production of relevant polyhydroxyalkanoates in microorganisms. *Biochem. Eng. J.* **16**, 81-96 (2003).

22. Sudesh, K., Bhubalan, K., Chuah, J-A., Kek, Y-K., Kamilah, H., Sridewi, N., Lee, Y-F. Synthesis of polyhydroxyalkanoate from palm oil and some new applications. *Appl. Microbiol. Biotechnol.* **89**, 1373-1386 (2011).

23. Anderson, J. & Dawes, E.a. Occurrence, metabolism, metabolic role, and

industrial uses of bacterial polyhydroxyalkanoates. *Microbiol. Rev.* **54,** 450–72 (1990).

24. Gao, D., Maehara, A., Yamane, T. & Ueda, S. Identification of the intracellular polyhydroxyalkanoate depolymerase gene of *Paracoccus denitrificans* and some properties of the gene product. *FEMS Microbiol. Lett.* **196,** 159–164 (2001).

25. Doi, Y., Kitamura, S., Abe, H. Microbial synthesis and characterization of poly(3-hydroxybutyrate-*co*-3-hydroxyhexanoate). *Macromolecules* **28**, 4822-4828 (1995).

26. Matsusaki, H., Abe, H., Doi, Y. Biosynthesis and properties of poly(3-hydroxybutyrate-*co*-3-hydroxyalkanoates) by recombinant strains of *Pseudomonas* sp. 61-3. *Biomacromol.* **1**, 17-22 (2000).

27. Sudesh, K. Polyhydroxyalkanoates from Palm Oil: Biodegradable Plastics. (Springer-Verlag Berlin Heidelberg) (2013).

28. Meng, D.-C., Shen, R., Yao, H., Chen, J-C., Wu, Q., Chen, G-Q. Engineering the diversity of polyesters. *Curr. Opin. Biotechnol.* **29,** 24–33 (2014).

29. Rehm, B. H. Biogenesis of microbial polyhydroxyalkanoate granules: a platform technology for the production of tailor-made bioparticles. *Curr. Issues Mol. Biol.* **9**, 41-62 (2007).

30. Sudesh, K., Abe, H., Doi, Y. Synthesis, structure and properties of polyhydroxyalkanoates: biological polyesters. *Prog. Polym. Sci.* **25**, 1503-1555 (2000).

31. Rehm, B.H.A. Polyester synthases: natural catalysts for plastics. *Biochem. J.* **376**, 15-33 (2003).

32. Schubert, P., Steinbüchel, A. & Schlegel, H.G. Cloning of the *Alcaligenes eutrophus* genes for synthesis of poly-beta-hydroxybutyric acid (PHB) and synthesis of PHB in *Escherichia coli*. *J. Bacteriol.* **170**, 5837–5847 (1988).

33. Peoples, O.P. & Sinskey, A. J. Poly-beta-hydroxybutyrate (PHB) biosynthesis in *Alcaligenes eutrophus* H16. Identification and characterization of the PHB polymerase gene (*phbC*). *J. Biol. Chem.* **264**, 15298–15303 (1989).

34. Liebergesell, M. and Steinbüchel, A. Cloning and nucleotide sequences of genes relevant for biosynthesis of poly(3-hydroxybutyric acid) in *Chromatium vinosum*

strain D. *Eur. J. Biochem.* **209**, 135-150 (1992).

35. Stubbe, J. and Tian, J. Polyhydroxyalkanoate (PHA) hemeostasis: the role of PHA synthase. *Nat. Prod. Rep.* **20**(5), 445-457 (2003).

36. Potter, M. & Steinbüchel, A. Poly(3-hydroxybutyrate) granule-associated proteins: impacts on poly(3-hydroxybutyrate) synthesis and degradation. *Biomacromolecules* **6**, 552–560 (2005).

37. Fukui, T. and Y. Doi. Cloning and analysis of the poly(3-hydroxybutyrate-co-3-hydroxyhexanoate) biosynthesis genes of *Aeromonas caviae*. *J. Bacteriol.* **179**, 4821-4830 (1997).

38. Jia, Y., Kappock, T.J., Frick, T., Sinskey, A.J. & Stubbe, J. Lipases provide a new mechanistic model for polyhydroxybutyrate (PHB) synthases: Characterization of the functional residues in *Chromatium vinosum* PHB synthase. *Biochemistry* **39,** 3927–3936 (2000).

39. Ollis, D. L. *et al.* The alpha/beta Hydrolase Fold. *Protein Eng* **5,** (1992).

40. Nardini, M. & Dijkstra, B. W. Alpha/beta hydrolase fold enzymes: the family keeps growing. *Curr. Opin. Struct. Biol.* **9,** 732–737 (1999).

41. Rehm, B.H.A., Antonio, R.V., Spiekermann, P., Amara, A. A. & Steinbüchel, A. Molecular characterization of the poly(3-hydroxybutyrate) (PHB) synthase from *Ralstonia eutropha*: In vitro evolution, site-specific mutagenesis and development of a PHB synthase protein model. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* **1594,** 178–190 (2002).

42. Amara, A. & Rehm, B.H.A. Replacement of the catalytic nucleophile cysteine-296 by serine in class II polyhydroxyalkanoate synthase from *Pseudomonas aeruginosa*-mediated synthesis of a new polyester: identification of catalytic residues. *Biochem. J.* **374,** 413–421 (2003).

43. Wahab, H. A., Ahmad Khairudin, N. B., Samian, M. R. & Najimudin, N. Sequence analysis and structure prediction of type II Pseudomonas sp. USM 4--55 PHA synthase and an insight into its catalytic mechanism. *BMC Struct. Biol.* **6,** 23 (2006).

44. Schubert, P., Kruger, N. & Steinbuchel, A. Molecular analysis of the Alcaligenes eutrophus poly(3-hydroxybutyrate) biosynthetic operon: identification of the N terminus of poly(3-hydroxybutyrate) synthase and identification of the promoter. *J. Bacteriol.* **173,** 168–175 (1991).

45. Zheng, Z. *et al.* Mutation on N-terminus of polyhydroxybutyrate synthase of Ralstonia eutropha enhanced PHB accumulation. *Appl. Microbiol. Biotechnol.* **72,** 896–905 (2006).

46. Tsuge, T. *et al.* Combination of N149S and D171G mutations in Aeromonas caviae polyhydroxyalkanoate synthase and impact on polyhydroxyalkanoate biosynthesis. *FEMS microbiology letters* **277,** 217–222 (2007).

47. Steinbuchel, A. *et al.* Molecular basis for biosynthesis and accumulation of polyhydroxyalkanoic acids in bacteria. *FEMS Microbiol. Rev.* **9,** 217–230 (1992).

48. Müh, U., Sinskey, A.J., Kirby, D.P., Lane, W.S. & Stubbe, J. PHA synthase from *Chromatium vinosum*: cysteine 149 is involved in covalent catalysis. *Biochemistry* **38,** 826–837 (1999).

49. Taguchi, S. & Doi, Y. Evolution of polyhydroxyalkanoate (PHA) production system by 'enzyme evolution': Successful case studies of directed evolution. *Macromol. Biosci.* **4,** 145–156 (2004).

50. Bhubalan, K., Kam, Y.C., Yong, K.H. & Sudesh, K. Cloning and expression of the PHA synthase gene from a locally isolated *Chromobacterium* sp. USM2. *Malays. J. Microbiol.* **6,** 81–90 (2010).

51. Bhubalan, K., Rathi, D.N., Abe, H., Iwata, T. & Sudesh, K. Improved synthesis of P(3HB-co-3HV-co-3HHx) terpolymers by mutant *Cupriavidus necator* using the PHA synthase gene of *Chromobacterium* sp. USM2 with high affinity towards 3HV. *Polym. Degrad. Stab.* **95,** 1436–1442 (2010).

52. Bhubalan, K. *et al.* Characterization of the highly active polyhydroxyalkanoate synthase of *Chromobacterium* sp. strain USM2. *Appl. Environ. Microbiol.* **77,** 2926–33 (2011).

53. Dong, A. *et al.* In situ proteolysis for protein crystallization and structure determination. *Nat. Methods* **4,** 1019–1021 (2007).

54. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods in Enzymol.* **276,** 307–326 (1997).

55. Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364,** 215–230 (2007).

56. Sheldrick, G.M. Experimental phasing with SHELXC/D/E: Combining chain tracing with density modification. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,**

479–485 (2010).

57. Abrahams, J.P. & Leslie, A.G.W. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **52,** 30–42 (1996).

58. Cowtan, K. The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **62,** 1002–1011 (2006).

59. Langer, G., Cohen, S.X., Lamzin, V.S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3,** 1171–9 (2008).

60. Murshudov, G.N. *et al.* REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67,** 355–367 (2011).

61. Adams, P.D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66,** 213–221 (2010).

62. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 2126–2132 (2004).

63. McCoy, A.J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

64. Kabsch, W. A solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A* **32,** 922–923 (1976).

65. Roussel, A. et al. Crystal structure of human gastric lipase and model of lysosomal acid lipase, two lipolytic enzymes of medical interest. J. Biol. Chem. 274, 16995–17002 (1999).

66. Roussel, A. et al. Crystal structure of the open form of dog gastric lipase in complex with a phosphonate inhibitor. J. Biol. Chem. 277, 2266–2274 (2002).

67. Jia, Y. *et al.* Mechanistic studies on class I polyhydroxybutyrate (PHB) synthase from *Ralstonia eutropha*: Class I and III synthases share a similar catalytic mechanism. *Biochemistry* **40,** 1011–1019 (2001).

68. Tian, J., Sinskey, A.J. & Stubbe, J. Detection of intermediates from the polymerization reaction catalyzed by a D302A mutant of class III polyhydroxyalkanoate (PHA) synthase. *Biochemistry* **44,** 1495–1503 (2005).

69. Wittenborn, E.C., Jost, M., Wei, Y., Stubbe, J. & Drennan, C. L. Structure of the Catalytic Domain of the Class I Polyhydroxybutyrate Synthase from *Cupriavidus necator*. *J. Biol. Chem.* **291,** 25264–25277 (2016).

70. Kim, J., Kim, Y.-J., Choi, S.Y., Lee, S.Y. & Kim, K.-J. Crystal structure of *Ralstonia eutropha* polyhydroxyalkanoate synthase C-terminal domain and reaction mechanisms. *Biotechnol. J.* **12,** 1600648 (2017).

71. Wodzinska, J. *et al.* Polyhydroxybutyrate synthase: Evidence for covalent catalysis. *J. Am. Chem. Soc.* **118,** 6319–6320 (1996).

72. Buckley, R.M. & Stubbe, J. Chemistry with an artificial primer of polyhydroxybutyrate synthase suggests a mechanism for chain termination. *Biochemistry* **54,** 2117–2125 (2015).

73. Ushimaru, K., Sangiambut, S., Thomson, N., Sivaniah, E., Tsuge, T. New insights into activation and substrate recognition of polyhydroxyalkanoate synthase from Ralstonia eutropha. *Appl. Microbiol. Biotechnol.* **97**(3), 1175-1182 (2013).

74. Gerngross, T.U., and Martin, D.P. Enzyme-catalyzed synthesis of poly[(*R*)-(-)-3-hydroxybutyrate]:Formation of macroscopic granules *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 6279-6283 (1995).

75. Zhang, W., Chen, C., Cao, R., Maurmann, L., Li, P. Inhibitors of polyhydroxyalkanoate (PHA) synthases: synthesis, molecular docking, and implications. *Chembiochem*. **16**(1), 156-166 (2015).

76. Stubbe, J. et al. Nontemplate-dependent polymerization processes: polyhydroxyalkanoate synthases as a paradigm. Annu. Rev. Biochem. 74, 433–480 (2005).

77. Gerngross, T.U. *et al.* Overexpression and purification of the soluble polyhydroxyalkanoate synthase from *Alcaligenes eutrophus*: evidence for a required posttranslational modification for catalytic activity. *Biochemistry* **33,** 9311–9320 (1994).

78. Griebel, R., Smith, Z. & Merrick, J. M. Metabolism of poly-beta-hydroxybutyrate. I. Purification, composition, and properties of native poly-beta-hydroxybutyrate granules from Bacillus megaterium. *Biochemistry* **7,** 3676–3681 (1968).

79. Chuah, J.A. *et al.* Characterization of site-specific mutations in a short-chain-length/medium-chain-length polyhydroxyalkanoate synthase: In vivo and in vitro studies of enzymatic activity and substrate specificity. *Appl. Environ. Microbiol.* **79,** 3813–3821 (2013).

80. Amara, A.A., Steinbuchel, A. & Rehm, B.H.A. In vivo evolution of the *Aeromonas punctata* polyhydroxyalkanoate (PHA) synthase: isolation and characterization of modified PH A synthases with enhanced activity. *Appl. Microbiol. Biotechnol.* **59,** 477–482 (2002).

81. Normi, Y.M. *et al.* Site-directed saturation mutagenesis at residue F420 and recombination with another beneficial mutation of *Ralstonia eutropha* polyhydroxyalkanoate synthase. *Biotechnol. Lett.* **27,** 705–712 (2005).

82. Taguchi, S., Nakamura, H., Hiraishi, T., Yamato, I. & Doi, Y. in Vitro Evolution of a Polyhydroxybutyrate Synthase by Intragenic Suppression-Type Mutagenesis. *J. Biochem.* **131,** 801–806 (2002).